An Introduction to Modern Theories of Economic Growth

8)

Hywel G. Jones

University Lecturer in the Economics of the Firm
and
Fellow of Linacre College, Oxford

Van Nostrand Reinhold (UK) Co. Ltd.

Contents

chapter 2
Concepts and Methods of Growth Theory

2.1 Introduction

The theory of economic growth is often alleged by students to be one of the more daunting specialized areas of an increasingly daunting subject. While there is undoubtedly some truth in the allegation, much of the difficulty stems from the plethora of esoteric symbols and concepts, contradictory assumptions and complicated techniques which surround ideas which are often really rather simple. Theories of growth have generated bitter controversy and, by their very nature, are particularly amenable to the application of difficult mathematical methods. As a consequence, although these theories have provided a major focus of professional interest in the past twenty years or so, the bewildered student is often deterred from attempting to gain any insight into the fascinating and important issues which are being discussed. Thus, a necessary preliminary to any study of the theories themselves is an investigation of some of the principal concepts, tools and methods which are employed in their elaboration. To the student who is eager to proceed to the theories, such a study may appear a rather tedious digression but a firm understanding of the concepts and methods will prevent, or minimize, confusion at a later stage. This chapter attempts to refresh the student's memory of some familiar concepts and introduces some unfamiliar ideas which will probably not have been encountered in the conventional elementary macroeconomics course.

2.2 The Variables and Aggregation

Theories of economic growth, particularly in their simpler forms, are usually conducted within the framework of a macroeconomic[1] model. Given Keynes' seminal role in the development of the macroeconomic method, it is hardly surprising to find that much of the nomenclature of the theory of macroeconomic growth stems from the host of concepts introduced in his *General Theory of Employment, Interest and Money* (140). Despite this similarity of terminology, we will see that the addition of older

[1] The term 'macroeconomics' was coined by Ragnar Frisch in 1933. It might be argued that some of the concepts associated with growth theory have been more fruitful in the sphere of microeconomics.

and newer concepts to his macrostatic model[2] can produce models of the long-run process of growth which abstract from, and even contradict, many of the central ideas in his own version of the moving forces of the economy as a whole.

This section discusses some of the principal aggregative variables which are the basic ingredients of most theories of economic growth. Other variables are introduced in the appropriate chapters as the book proceeds.

(a) National Income or Product

The national output of an economy is, of course, the variety of goods and services produced in the economy during any particular period of time. As it is necessary to specify the time period—e.g. national output per year or per quarter—it should be clear that this is a flow concept[3]. Since there is no obvious way in which the output of the steel industry can be added to the 'output' of a hairdresser or a university lecturer, it is usual to aggregate (i.e. 'add-up') all the heterogeneous outputs in the economy by measuring each output in *value* terms—i.e. by multiplying the output of any industry by an appropriate price—usually the market price of the good or service. The student will be familiar with the idea that, with appropriate adjustments, the resulting total will be equal to the sum of the *incomes* of all firms and households or to the sum of the *expenditures* of all firms and households.

It should be clear that national product, as measured by a government statistical office, can vary over time despite the absence of any change in the underlying real output produced by the economy. This will, of course, be the case when prices are changing and the student will undoubtedly be familiar with the difference between national income estimates measured in *current* as opposed to *constant* prices. Moreover, it is necessary to distinguish between the *actual* and the *potential* level of income. At any point in time a certain level of income *could* be generated if the economy was fully employing all the factors available—i.e. the economy was on the *production possibility frontier* (see Samuelson (227) p. 22). But it might be that the *actual* level of income is well below the maximum potential level of income. In this case a distinction should be drawn between growth in actual income and growth in maximum potential income.

Although it is clear that national income estimates are an important requirement of the practical economic policy maker, it should be noted that there are serious conceptual problems involved in using such a construct in the context of a theoretical analysis. Keynes, for example, although associated in many students' minds with the development of the

[2] The 'new' view of Keynes' contribution (see (158) (159) and (110)) would suggest that his overall 'vision' of the macroeconomy was not as static as has been suggested in the literature.

[3] See, for example, Samuelson (227) Ch. 10; Lipsey (164) Ch. 34 or Brooman (29) Ch. 2.

idea of national income, protested vigorously against the use of 'incommensurable collections of miscellaneous objects' ((140) p. 39) for the purposes of causal analysis. We will return to these problems later in the section. (See 2.2(d).)

Throughout this book we will use the symbol Y to refer to national income; where there is need to distinguish between income and output we will refer to national output as Q.

(b) The Stock of Capital

Few concepts in economics have been subject to so many different interpretations and provoked so much controversy over so long a period of time as that of *the* stock of 'capital'. Let us consider some of the different interpretations that are available.

(i) To the economist interested in economic growth, 'capital' is generally taken to mean the *stock of produced means of production available to a firm or an economy at any point in time*—i.e. the stock of fixed capital equipment—although most simple models of growth subsume *land* within the general concept of capital and in some contexts, it is useful to think of a firm's capital as including inventories and work-in-progress. Such an interpretation is valid whatever the institutional setting—it is as true for the USSR as for the USA. However, in a 'capitalist' economy, 'capital' is *property* and conveys rights to a share (profit) of the product to its private owners—which is not, of course, a characteristic of economic life in socialist countries. If 'capital' is viewed as a stock of physical objects—machines, factories, etc.—then *Net Investment* is the flow of new machines which augment this stock[4].

(ii) In recent years economists have increasingly used special concepts of capital and investment in rather different contexts to those specified in (i) above. '*Social overhead capital*' refers to the stock of roads, bridges, ports, etc. which, although not contributing directly to the production of output, provide the essential framework in which conventional economic activity can take place, and facilitate investment. Educational spending is often referred to in modern discussions as 'investment' in '*human capital*' and it should be clear that this usage is very different to the conventional. It is sometimes useful to distinguish between capital goods which are used for producing more capital goods and those which are used for the production of final goods for consumption. (See 5.6.)

(iii) The businessman often uses the terms 'capital' and 'investment' in a rather different way from those specified above. The term 'capital' is sometimes used to mean a sum of money available for investment in the

[4] Gross Investment refers to the *total* of new capital created during a period of time—some part of which is used for the replacement of old machines. Thus, Net Investment = Gross Investment − Depreciation. See Brooman (29) Ch. 2.

expectation of profit and sometimes the actual plant and machinery already in use. The term 'investment' is even more liable to cause confusion. In particular, it should be noted that the purchase of shares in a company from a stock exchange does not usually[5] imply a net addition to the stock of capital of the firm or the economy as a whole—and is therefore *not* investment in the economists' sense as used in this book.

Most simple models of economic growth confine themselves to the concepts of capital and investment specified in (i) above and, for the most part, we will follow this procedure throughout this book.

Consider then, a firm—call it firm 1—which produces an output or outputs using a great variety of different items of capital equipment, of different ages, and for different specific purposes. Number these different machines from 1 to n and the capital equipment of the firm can be *listed* as follows:

$$(K_{11}, K_{12}, K_{13} \ldots K_{1n}) \tag{2.2.1}$$

where K_{11} refers to the amount of equipment of type 1 used by firm 1, K_{12} the amount of equipment of type 2 used by firm 1 and so on[6] (the first subscript refers to the firm and the second to the type of capital equipment). It should be clear that the list or vector[7] of capital goods used by the firm will be very long for all but the simplest enterprise. Now, suppose that there are a large number of firms—say m—in the economy. Each of the m firms will now have its own list or vector of capital goods and the capital stock of the country, excluding any concept of social capital, can be represented by a rectangular array of numbers as follows:

$$\begin{matrix} K_{11}, & K_{12}, & K_{13}, & K_{14} \ldots & K_{1n} \\ K_{21}, & K_{22}, & K_{23}, & K_{24} \ldots & K_{2n} \\ K_{31}, & K_{32}, & K_{33}, & K_{34} \ldots & K_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ K_{m1}, & K_{m2}, & K_{m3}, & K_{m4} \ldots & K_{mn} \end{matrix} \tag{2.2.2}$$

where K_{32} is the symbol for the amount of capital equipment of type 2 used by firm 3 and K_{mn} is the amount of capital equipment of type n used by firm m. If, as would presumably be the normal case, a particular

[5] The purchase of shares on a stock exchange usually involves no more than the transfer of a claim on a company from one person to another. Only in the case of a *new issue* by a firm is additional finance for real investment made available to it by means of a stock exchange. See Brooman (29) p. 41.

[6] The series of dots, . . . , in expression (2.2.1) represents all the types of equipment from type 3 to type n.

[7] 'Vector' is a mathematical expression for an ordered array, or list, of numbers. For more detail, see Mills (181) Ch. 2.

firm does not use every item of capital equipment then the corresponding entry in the array will, of course, equal zero.

Is there any way whereby this representation of the capital stock can be simplified so that it is possible to speak of *the* capital stock as a single aggregate? It might seem that one simplifying method would be the addition of all the goods of each type—i.e. those that are technically identical so as to produce a single vector of capital goods for the economy as a whole:

$$[K_1, K_2 \ldots K_n] \tag{2.2.3}$$

where K_i is the total number of technically identical machines of type i used by all the m firms and $i = 1, 2 \ldots n$[8]. However, it should be noticed that even if the machines of type i used by firm 1 are *technically* identical to those of type i used by firm 2 they may differ in economic significance in terms of the flows of returns, and, in the capitalist sense, profits, that they can be expected to earn during the respective lifetimes, e.g. two technically identical ice-cream machines clearly differ considerably in economic significance if one is sited at Blackpool or Coney Island and the other in Warrington or Toledo, Ohio!

If, for the sake of argument, it was accepted that capital goods of each type were *economically* and *technically identical* such that the vector of capital goods (2.2.3) was an acceptable simplification of the array of (2.2.2) and Gertrude Stein's dictum that 'a spade is a spade is a spade'[9] was upheld, we would still face something of an impasse. Each of the n heterogeneous capital goods is fundamentally different from each other. It is, in general, *specific* to the task for which it was designed—i.e. a tractor, say, cannot be used to produce car bodies without some Heath–Robinson-like improvisation! The problem of producing a measure of *the* capital stock as a whole has produced one of the most vigorous and long-standing debates in the history of economic thought and this is given further consideration in Section 6.2. For the moment, we note that any simple aggregate measure of the heterogeneous capital stock would require a common standard into which all the different items of capital equipment could be converted and then added together. The conventional approach would be to value 'capital'. But should 'capital' be valued by the costs incurred (in the past) in its production or through the future flows of earnings that can be expected to flow from its usage? Eschewing controversy for the moment we note that a conventional method of valuing 'capital' is by calculating the PRESENT VALUE (see, for example, Brooman (29) pp. 150–4) of the expected returns associated with each item of capital stock.

[8] (2.2.3) is derived from (2.2.2) by adding the elements in each column to produce totals of each type of 'machine' in the economy.

[9] Actually Stein wrote: 'Rose is a rose is a rose is a rose' (*Sacred Emily*).

Consider the simple compound interest problem of elementary algebra. If a sum Z is invested for t years at compound interest it will, at the end of the period, amount to a sum A where

$$A = Z(1 + r)^t$$

and r is the rate of interest.
If we wish to calculate the original sum from a knowledge of A, r and t we simple rearrange the formula as

$$Z = \frac{A}{(1 + r)^t}$$

Now, if a firm buys a machine X, *expecting* a lump-sum return of A, t years hence, it is clear that the current value of the machine (usually referred to as the PRESENT VALUE) is given by

$$\text{Present value of X} = \frac{A}{(1 + r)^t} \tag{2.2.4}$$

The returns from an item of capital equipment typically accrue over a number of years. Consider, for example, a farmer who owns a tractor. He anticipates that it will last for another five years before needing replacement. He expects returns A_1, A_2, A_3, A_4 and A_5 during the next five years. Then the present value is given by

$$\begin{aligned}\text{Present Value of Tractor} = {} & \frac{A_1}{(1 + r)} + \frac{A_2}{(1 + r)^2} + \frac{A_3}{(1 + r)^3} + \frac{A_4}{(1 + r)^4} \\ & + \frac{A^5}{(1 + r)^5}\end{aligned} \tag{2.2.5}$$

Now, if all the capital goods in the economy are valued in this way (i.e. all the items in the array (2.2.2) then the respective present values can be added together to produce a single index of the capital stock measured in terms of *value capital*. Notice that the resulting aggregate depends upon the rate of interest at which the future expected flows are discounted, and the expected future flows themselves, which depend upon the present and future relative prices of the outputs produced by the items of capital equipment.

We have dwelt for some time on the concept of capital partly because it is central to any theory of growth and partly because it is essential to perceive at an early stage the difficulties of talking of aggregate 'capital'. The symbol K will be used throughout this book when referring to an index of aggregate capital.

(c) The Stock of Labour

We need not dwell for long on the concept of the stock of labour. It is clear that, in any real economy, there are a great variety of different kinds of labour of differing skills, intrinsic ability and experience. Simple models of growth tend to assume an *homogeneous* labour force and measure the stock in terms of the number of labourers and the flow of labour services in terms of man-hours. Moreover, the labour force is generally assumed to be a fixed proportion of the total population: $L = aP$ where L = Labour force, P = Population and a is a constant. This assumption implies that, in the case of a growing economy, the labour force L, grows at the same rate as the whole population P. Simple growth models usually assume that the population, and hence the labour force, grows at a *constant*, *exogenous* rate. This means that no element of the economic model under consideration can affect the rate of population growth and Malthusian mechanisms, for example, are assumed away[10].

The above assumptions are clearly drastic simplifications. Real growing economies are often constrained by shortages of particular kinds of skilled labour, by a declining proportion of the population entering the labour force as fewer women and children work or by a low average educational level of the labour force as a whole. One would expect that the proportion of the total population that worked would depend on the relative attractiveness of working, i.e. on the relative valuations of work and leisure by the worker. The rate of growth of the population might be expected to depend upon the average level of income of each household. However, for our purposes, the simplest assumption will be seen to be reasonably fruitful and we will discuss the effects upon our conclusions of introducing some of the complications suggested above.

(d) Aggregation and Parable

The preceding sections have highlighted some of the difficulties associated with the central concepts of an *aggregative* model of growth. Yet some degree of aggregation is necessary as is clear from the following comments from one of the most persistent and resourceful *critics* of the concepts of aggregate 'capital'.

> 'A model which took account of all the variegation of reality would be of no more use than a map at the scale of one to one'

although

> 'We must be careful not to make a simplification in such a way that the model falls to pieces when it is removed'

[10] On Malthus' theory of population, see Blaug (25) Ch. 3.

and

> 'A highly aggregated model is useful only for a first sketch of the analysis of reality, but it is much easier to fill in the details of the outline drawn by a simple model than it is to build up an outline by assessing details' (Professor Mrs Joan Robinson (212) pp. 33 and 34).

In simple models of growth, aggregative concepts are often used with little attempt at justification—a procedure which is often referred to as 'heroic' aggregation. On the other hand, it is possible to try and construct a *rigorous* aggregative model in which the conclusions of the simplified construct can be shown to be identical with those that are derived from a detailed model which eschews the use of aggregative concepts. (See 6.2.)

One approach to aggregation difficulties is to conduct one's theory of growth solely in terms of 'parable'. To quote Solow:

> 'We are dealing with a drastically simplified story, a "parable" which my dictionary defines as "a fictitious narrative or allegory (usually something that might naturally occur) by which moral or spiritual relations are typically set forth". If moral or spiritual relations, why not economic?' ((252) p. 1).

The most common form of this 'parable' is the assumption that only a *single* good—call it 'corn'—is produced in the economy. In this case the index number difficulties of sections (a) and (b) are totally removed. The output of the economy is unambiguously defined in terms of tons of 'corn'. The 'corn' is either eaten (consumed) or 'invested'—in which case it becomes part of the 'corn capital' stock. This 'parable' approach to analysing economic growth would, of course, be completely valid and useful if it could be shown that it did not, in some cases, distort the conclusions of the analysis despite the drastic simplification. This problem will recur throughout the book but particularly in Chapter 6.

(e) Equilibrium Conditions and Accounting Identities

The distinction between equilibrium conditions and accounting identities[11] is often crucial for the correct interpretation of models of economic growth. (See, for example, Ch. 3.) The distinction is particularly easy to appreciate in the context of the elementary theory of the determination of price in a competitive market. In any market, the quantity *actually supplied* in any period of time *must* equal the quantity *actually purchased*—for to every purchase there corresponds a sale. The apparatus of supply and demand curves, however, refers to *plans* to buy and sell at different prices—the *equilibrium* price being determined by the equality of *planned supply*

[11] This distinction is extensively discussed in most elementary textbooks. See, for example, Lipsey (164) Chs 7–9, Rowan (218) Chs 3 and 4 or Brooman (29) Ch. 3.

and *planned demand*. Thus, if we write D = Demand and S = Supply, and use the subscript A to indicate actual values, the statement

$$D_A \equiv S_A \tag{2.2.6}$$

is fundamentally different from the statement

$$D = S \tag{2.2.7}$$

(2.2.6) expresses the fact that, in an accounting sense, actual supply and demand *must* be equal whereas (2.2.7) is the equilibrium condition for a competitive market. In the context of a macroeconomic model, *actual investment must* equal *actual* saving as a simple consequence of the conventional method of national accounting. (See (218) pp. 64–7.) On the other hand, *planned* investment need not equal *planned* savings: the intersection of the schedules of planned saving and planned investment determines the level of *equilibrium* national income in a simple Keynesian macroeconomic model. Some writers use the phrases 'ex-ante' and 'ex-post' to distinguish between planned and actual quantities. Thus, in this terminology, ex-post saving *must* equal ex-post investment whereas the equality of ex-ante saving and investment is the equilibrium condition of a simple macroeconomic model.

These distinctions appear and reappear in various guises throughout this book: if the reader is not confident with them he should consult any of the various cited texts.

2.3 Saving and Investment

The determinants and saving and investment play a crucial role in many simple models of economic growth and it is therefore necessary to review these concepts although they may already be familiar to the student from elementary macroeconomics.

(a) The Savings Function

Simple models of economic growth, following the Keynesian tradition, invariably assume that the aggregate savings in an economy are a simple function of (i.e. 'depend upon') the level of income Y in the economy: $S = S(Y)$ where S = aggregate saving and Y = the level of income. Despite the emergence of more sophisticated[12] theories of saving from Friedman (80) and Modigliani, Brumberg and Ando (185) in particular, many simple theories of growth assume that aggregate savings are proportional to aggregate income:

$$S = sY \tag{2.3.1}$$

where s corresponds to both the average and marginal propensities to save. The marginal (and average) propensity to save is usually assumed to be positive but less than one ($0 < s < 1$) which means, of course, that a part, but not the whole, of any increment in income is saved. Such a function is easily graphed as in Fig. 2.1 where the slope of the line is equal to s.

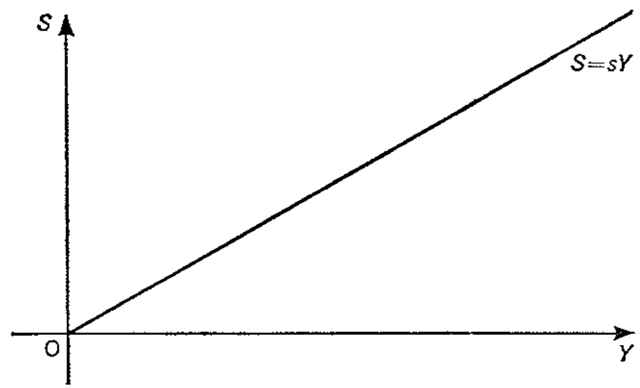

Fig. 2.1

Although advanced growth models incorporate more specialized savings mechanisms, this simple formulation is remarkedly fruitful in highlighting some of the relationships between saving and economic growth. Another formulation of the saving functions which is frequently used in the context of models of economic growth involves drawing a distinction between wage income and profit income. In this version, the aggregate savings function takes the form

$$S = s_w W + s_p P \quad \text{with} \quad Y \equiv W + P \quad \text{and} \quad 0 < s_w < s_p < 1 \tag{2.3.2}$$

where
W = Wage Income, P = Profits
s_w = propensity to save out of wages
s_p = propensity to save out of profits

A particular form of this function known as the *classical savings function* and often attributed to Kalecki (see (126)) involves the assumption of a zero marginal propensity to save out of wages:

$$S = s_p P \quad \text{with } 0 < s_p < 1 \tag{2.3.3}$$

[12] A thorough reading of Chs 8–10 of Keynes (140) suggests that many of the elements of the more sophisticated theories were certainly considered by him, if not in the modern form of attempting to derive macroeconomic savings behaviour from the study of the microeconomics of the 'rational' saver. Evans (67) Chs 2 and 3 contains a useful summary of the 'new' theories.

It is clear that in both these cases the overall propensity to save of the economy depends upon the *distribution of income* between wages and profits.

(b) Investment

Two concepts related to aggregate investment can be important in the context of models of economic growth—the Keynesian concept of the *multiplier* and the older concept of the *accelerator*—which is a particular case of the general idea of an INVESTMENT FUNCTION.

(i) THE MULTIPLIER

Readers will be familiar with the elementary idea of the multiplier whereby an increment in investment produces a greater increment in aggregate output. This simple idea, described by some as a 'mechanical toy', is usually attributed to R. F. Kahn (now Lord Kahn) (124), a student of Keynes. The idea, as opposed to its formalization, had been implicit in a polemical pamphlet that Keynes had published with H. D. Henderson (138) and he had used the concept in numerous articles prior to the publication of *The General Theory* (140). It was, however, its usage in *The General Theory* that began its progress to its current hallowed position in the textbooks. (See Samuelson (227) Ch. 12, or Brooman (29) Ch. 6.) There are many ways of demonstrating the fundamental point. Equilibrium in a closed economy with no government expenditure or foreign trade necessitates the equality of *planned* (i.e. ex-ante) saving and investment

$$I = S \tag{2.3.4}$$

Given the simple proportional savings function, this implies

$$I = sY \tag{2.3.5}$$

Consider an increment ΔI in I. This generates any increment ΔY in Y:

$$\Delta I = s\Delta Y \tag{2.3.6}$$

Hence, the ratio of the increment in Y to the increment in I is given by $1/s$. Since $0 < s < 1$ this implies that the increment in I generates a greater increment in Y. The student unfamiliar with this proposition should consult the appropriate section of the structured reading list.

(ii) INVESTMENT FUNCTIONS AND THE ACCELERATOR PRINCIPLE

A central feature of the Keynesian analysis is the observation, which has caused much controversy, that although S and I *must* be identical *ex-post*, savings and investment decisions are, in general, taken by different decision-makers and there is no reason why ex-ante savings (assumed to be a function of income) should equal ex-ante investment. If we accept this position then it is clearly necessary to analyse the determinants of investment separately from those of saving. It is, on the other hand, possible, in an alternative framework, to ignore discrepancies between ex-ante saving and investment by either:

(1) constructing a model in which ex-ante saving *is* identically investment—as in some of the simplified 'parables' of growth or

(2) assuming that some agency (the government?) uses the policy instruments at its disposal so as to maintain ex-ante I equal to ex-ante S. (See Ch. 4.)

In either of the above cases an *independent* investment function is not necessary and we will study models relying on the above assumptions in later chapters. (See, in particular, Ch. 4.)

Given that many models of growth, particularly those of Keynesian origin, *do* utilize an independent investment function, it is clearly necessary that we briefly examine some of the possibilities before proceeding.

The *acceleration principle*[13] is based upon the simple idea that the capital stock desired, or considered appropriate, by entrepreneurs as a whole depends upon the *level of demand* for the aggregate output that they produce which ultimately, of course, is given by the level of national income Y. In this case, net investment, which is an *increase* in the stock of capital, is related to the *increase* in the level of national income.

Thus, at its simplest, the accelerator principle can be written as

$$I_t = v[Y_t - Y_{t-1}] \tag{2.3.7}$$

i.e. aggregate net investment at time t, I_t, is a fixed proportion, v (the accelerator coefficient), of the difference between the demand for output in period t, Y_t, and the demand for output in the previous period, Y_{t-1}.

It should be clear, however, that entrepreneurs would normally be forced to make their investment decisions on the basis of their *expectations* of the demand for output. Another simple formulation of the acceleration principle which captures this idea can be written as

$$I_t = v[Y_t^E - Y_{t-1}] \tag{2.3.8}$$

i.e. net investment at the beginning of period t is a fixed proportion v of the difference between the level of demand for output in the last period, Y_{t-1}, and the level of demand for output *expected* in the coming period, Y_t^E.

There are serious deficiencies in the concepts of the accelerator, some of which are outlined in Matthews (175) Ch. 3, and the *capital stock adjustment principle* (C.S.A.) was an attempt to produce a different formu-

[13] The idea originated with Clark (44). For a simple exposition see Matthews (175) Ch. 3 or Brooman (29) pp. 168–73.

lation of the accelerator while retaining the central idea. The C.S.A. suggests that the volume of investment varies directly with the *level* of national income and inversely with stock of capital already in existence. Thus a linear version, incorporating expectations, of this principle can be written

$$I_t = aY_t^E - bK_t \tag{2.3.9}$$

where the symbols have an obvious meaning. The student should ask himself whether either of the above principles are 'rational' for the individual entrepreneur.

An alternative approach to the analysis of the determination of the level of aggregate investment stems from the idea of *present value* revised in Section 2.2(b). The student should be familiar with the idea that this approach, which suggests that individual investments will be made if their present value exceeds their present cost, or the alternative *internal rate of return* approach, both lead to the conclusion that the level of aggregate net investment will be a function of the rate of interest[14] and that expectations will again have a crucial role to play. Since this approach will not be particularly important in any of the models of growth that we analyse we will not go into detail although we will be free to discuss the ways in which the introduction of such an investment function would alter the conclusions of our analyses.

2.4 The Technology of the Economy

It is clear that the quantity of output produced by any economy is constrained by the available supplies of capital and labour[15]. It seems difficult to deny Samuelson's dramatic assertion:

> 'Until the laws of thermodynamics are repealed I shall continue to relate outputs to inputs' ((223) p. 444).

This kind of relationship is often summarized in an *aggregate production function*

$$Y = F(K, L) \tag{2.4.1}$$

which states that aggregate output, Y, is a function of the amounts of capital, K, and labour, L, in the economy. Following the discussion of the difficulties and controversy associated with the aggregative variables used in (2.4.1), readers will not be surprised to learn that the use of an aggregative production function, particularly in some of the specialized forms discussed below, is especially controversial.

[14] See, for example, Brooman (29), Ch. 7, Junankar (123), or, at a higher level, Lund (167).

[15] See Samuelson (227) Ch. 2. For a thorough mathematical discussion of most of the topics of 2.4 see Allen (7) Ch. 3.

The production function (2.4.1) is usually interpreted as indicating the *maximum* flow of output associated with the given amounts of capital and labour. K and L are sometimes interpreted as stocks and sometimes as flows of capital and labour services respectively. A careful reading is often required in order to ascertain which interpretation is being used in any given context. Two forms of the general aggregate production function of (2.4.1) are of interest to us: the *fixed coefficient form* and the *continuous form*.

(a) Fixed Coefficients

This simple form of aggregate production relationship has output, Y, determined in direct proportion to the quantities of capital and labour. Thus, $Y = K/v$ or L/u where v and u are constants. Different interpretations can be given to these fixed coefficients and some are discussed in Section 3.6. This form of production relationship basically implies that, given any particular stock of capital, there is one and only one flow of output which can be generated—and similarly for any given stock of labour. The actual production function takes the form

$$Y = \min\left[\frac{K}{v}, \frac{L}{u}\right] \tag{2.4.2}$$

(where min indicates minimum)

Say, for example, that L/u is the minimum of (2.4.2.) then Y is determined by L/u and the capital requirement by vY. This implies, of course, that the formulation (2.4.2) allows *either* capital or labour to remain *unused*. This form of technology implies that there is *no substitution* between capital and labour in the production of output. Given a certain quantity of K, one and only one flow of Y can be produced *however* much more labour is available. This form of the aggregate production function can be graphed as in Fig. 2.2, where the point $\bar{Y}$ indicates the one and only combination of K and L that can be used to produce a given amount of output, $\bar{Y}$. If more labour is available, say L^+, then the amount u_1L^+ is redundant and will remain unused. Similarly, if more capital, say K^+, is available then the amount v_1K^+ of capital is redundant. More or less output is obtainable only by *radial expansion or contraction* along the ray OZ—along which the amounts of capital and labour are kept strictly in proportion. (See Allen (7) pp. 35–7.)

(b) The Continuous Aggregate Production Function

(i) INTRODUCTORY

This form allows for the substitution of aggregate capital for labour in the production of output. Thus, any given flow of output $\bar{Y}$ can be produced

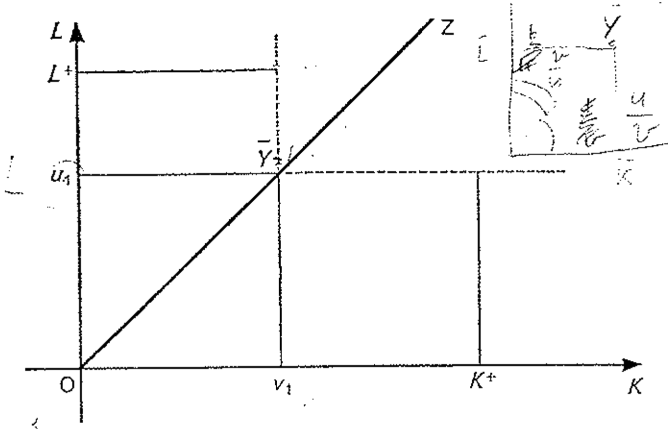

Fig. 2.2

by a *variety* of *combinations* of capital and labour. With continuous substitution possibilities, this form can be illustrated as in Fig. 2.3. In Fig. 2.3 the curve AB, which is called an 'isoquant' in a microeconomic context, indicates the different possible combinations of aggregate capital and labour that can produce the fixed flow of national income or output $\bar{Y}$. Thus, point A, involving a large amount of labour and a small amount of capital, is equivalent to point B, involving a relatively small amount of labour and a large amount of capital, in the production of the level of output $\bar{Y}$.

(ii) THE MARGINAL PRODUCTS OF CAPITAL AND LABOUR

With the continuous form of the aggregate production function it is possible to discuss the effect upon total output of a marginal increment in either capital or labour. We define the marginal product of labour as the

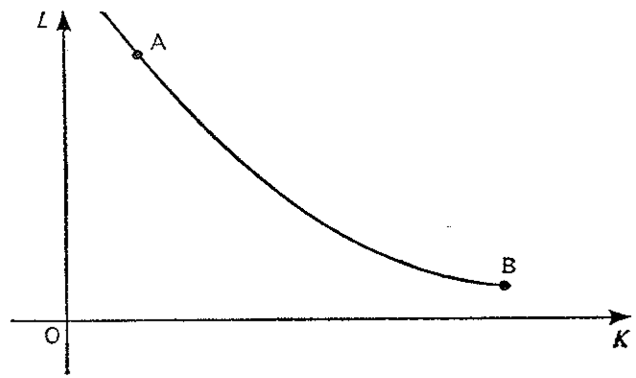

Fig. 2.3

extra output generated by an increment in the labour force (or supply of labour services), the capital stock being held constant. Rigorously, the marginal product of labour is the rate of change of output, Y, with respect to a change in labour, L, and is written $\partial Y/\partial L$ i.e. the partial derivative of output with respect to labour[16]. The marginal product of capital $\partial Y/\partial K$, is defined in a completely analogous fashion as the rate of change of output with respect to a change in K, the stock of labour being held constant.

Given our definition of marginal products we can further specialize the continuous form by assuming:

ASSUMPTION 2.4.1 POSITIVE MARGINAL PRODUCTS

The marginal products of capital and labour are both positive, i.e.

$$\frac{\partial Y}{\partial K} > 0 \text{ and } \frac{\partial Y}{\partial L} > 0$$

This means, of course, that an increase in either capital or labour will *always* increase the flow of output. Elementary microeconomics often assumes the possibility of the marginal products becoming eventually negative as more and more of one factor is added to a fixed factor. Assumption 2.4.1 rules out this possibility.

The student of elementary economics will undoubtedly be familiar with the following assumption (See Samuelson (237) pp. 24–7 or Lipsey (164) p. 215) which further specializes the continuous form of the production function:

ASSUMPTION 2.4.2 DIMINISHING MARGINAL PRODUCTIVITY

Although each increment in capital or labour generates an increment in the flow of output, successive *increments in capital or labour produce* decreasing increments *in the flow of output*. This assumption can be written rigorously as

$$\frac{\partial^2 Y}{\partial K^2} < 0 \text{ and } \frac{\partial^2 Y}{\partial L^2} < 0 \qquad \text{i.e. the rate of change of the}$$

rate of change of output with respect to a change in capital or labour is negative. This assumption corresponds to the familiar idea of Diminishing Returns to either factor with the other being held constant but, in elementary microeconomics, diminishing returns are not usually assumed to set in

[16] The non-mathematician should not despair. $\partial Y/\partial L$ should simply be read as 'the rate of change of Y with respect to a change in L with K held constant'. However, a small investment in reading Allen (6) Ch. VI and Ch. XI, or the relevant chapters in other textbooks such as Casson (37) or Archibald and Lipsey (8), would pay substantial dividends.

immediately. Assumption 2.4.2 implies that diminishing returns are present over the *whole range* of output.

(iii) LINEAR HOMOGENEITY OR CONSTANT RETURNS TO SCALE
A production function is said to be linearly homogeneous, i.e. it operates subject to constant returns to scale, if the multiplication of *both* capital and labour by a positive number implies that the output generated is multiplied by the same number

$$F(\lambda K, \lambda L) = \lambda F(K, L) = \lambda Y \text{ all } \lambda > 0 \tag{2.4.3}$$

i.e., if we double the amounts of capital and labour in the economy then the flow of output is doubled. Continuous aggregate production functions are often assumed to be of this form, so we note for future reference:

ASSUMPTION 2.4.3 CONSTANT RETURNS TO SCALE

The aggregate production function is linearly homogeneous.

It should be clear that the assumption of constant returns to scale is not incompatible with Assumption 2.4.2. The 'law' of diminishing returns refers to a situation where the marginal productivity of a factor decreases as the quantity of the factor employed is increased *with the quantities of other factors being held constant*. Constant returns to scale are defined in the case where *all* factors are increased in the same proportion.

The assumption of constant returns to scale permits a substantial simplification of the aggregate production function in that it can be written in a *per-worker* or '*intensive*' form. Given a constant returns aggregate production function, $Y = F(K, L)$, we know that multiplying both K and L by some number λ will result in Y being multiplied by the same number. Put $\lambda = 1/L$ and multiply through to obtain

$$\frac{Y}{L} = F\left[\frac{K}{L}, 1\right] \tag{2.4.4}$$

Equation (2.4.4) simply states that outputs per worker, Y/L, depends upon capital per worker or the capital–labour ratio, K/L. (2.4.4) can be simply written as

$$y = f(k) \tag{2.4.5}$$

where $\quad y = \dfrac{K}{L},\ k = \dfrac{K}{L}$ and $f(k) = F(k, 1)$

Equation (2.4.5) is the per-worker form of the aggregate production function and it constitutes a central tool of many models of economic growth. It is used very frequently in this book. If we make one further

assumption then the per-worker production function can be easily illustrated in a diagram.

ASSUMPTION 2.4.4 NO INPUT, NO OUTPUT
If no capital or labour is employed then no output can be produced, i.e. given $y = f(k)$ if $k = 0$ then $y = 0$.

Given assumptions 2.4.1–2.4.4, the aggregate production function can be illustrated as in Fig. 2.4. Each point on the curve $f(k)$ in Fig. 2.4 shows

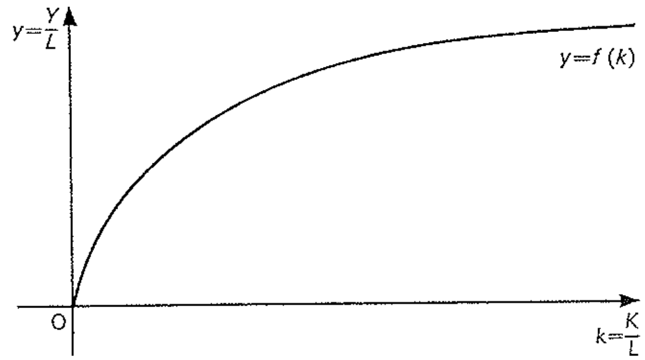

Fig. 2.4

the quantity of output per worker produced by the corresponding input of capital per worker. If the labour force is held constant and, by a suitable choice of units, put equal to one, then Fig. 2.4 can be considered as illustrating the relationship between total product and aggregate capital employed. Assumption 2.4.4 ensures that the curve begins at the origin. Assumption 2.4.1 implies that it is upward sloping and Assumption 2.4.2 is the justification for its 'flattening out'. As might be expected, it is particularly convenient that the production possibilities or technology of a model economy can be so simply illustrated and the curve $f(k)$ appears as a constituent part of many of the other diagrams in this book.

(iv) THE MARGINAL PRODUCTIVITY THEORY OF DISTRIBUTION
A conventional result of microeconomic theory demonstrates that a profit-maximizing entrepreneur will hire factors of production (e.g. capital or labour) up to the point where their marginal revenue products equal their price. (See Lipsey (164) Part 6 or Samuelson (227) Part 4.) When translated into the context of the macroeconomy this theory, under the title of the 'marginal productivity theory of distribution', suggests that, in competitive conditions, the real wage rate of labour will be equated to 'the' marginal product of labour as a whole and that the real rental for a unit of capital will be equated to 'the' marginal product of capital as a whole.

The theory was developed by the so-called neoclassical theorists (see

4.1) in the last quarter of the nineteenth century. A series of writers—J. B. Clark in the United States, Marshall and Wicksteed in England, Walras in Switzerland, Wicksell in Sweden and Barone in Italy all produced theories which, with remarkable simultaneity, incorporated the substance of marginal productivity theory[17]. The intense controversy that has surrounded this theory for the eighty or so years of its existence (see Ch. 6) stems not only from the theoretical problems associated with it but also from the attempts of some of its proponents, notably J. B. Clark, to indulge in what Stigler has called 'naive productivity ethics', i.e. to imply that free competition produces a 'just' wage and distribution of income. In Clark's words:

> 'Where natural laws have their way, the share of income that attaches to any productive function is gauged by the actual product of it. In other words, free competition tends to give to labour what labour creates'[18] (J. B. Clark (42) p. 3).

Given the subsequent controversies (see Ch. 6), it is ironic to reflect that the preface to Clarke's *Philosophy of Wealth* outlined the belief that 'the period of irreconcilable diversity in the fundamental principles of the science seems past and an era of relative unanimity . . . appears to have arrived'.

If capital and labour are paid their marginal products what guarantee is there that the total product will be exactly exhausted—i.e. there will be no surplus or shortfall? A mathematical theorem due to *Euler* guarantees that if the production function is subject to constant returns to scale then the payment of marginal products to the factors of production will exactly exhaust the product (see Allen (6) pp. 317–19)

i.e. Quantity of Capital × Marginal Product of Capital + Quantity of Labour × Marginal Product of Labour = Total Output

or, in mathematical form,

$$K\frac{\partial Y}{\partial K} + L\frac{\partial Y}{\partial L} = Y \tag{2.4.6}$$

where the meaning of $\partial Y/\partial K$ and $\partial Y/\partial L$ is explained in Section 2.4(ii). If the marginal productivity theory of distribution is accepted then the price of capital, the real profit rate, is equated to the marginal product of capital

[17] For a discussion of the historical background of the theory see Stigler (257) or Blaug (25) Ch. 11.

[18] A thorough reading of Clark (42) is required to get the full flavour of his approach.

and the price of labour, the real wage rate, is equal to the marginal product of labour.

The distribution of income between capital and labour can be simply illustrated in the context of the diagram of Fig. 2.4. Consider Fig. 2.5. This appears identical to Fig. 2.4 but the student will notice that we have explicitly assumed the labour force constant at unity so that the national

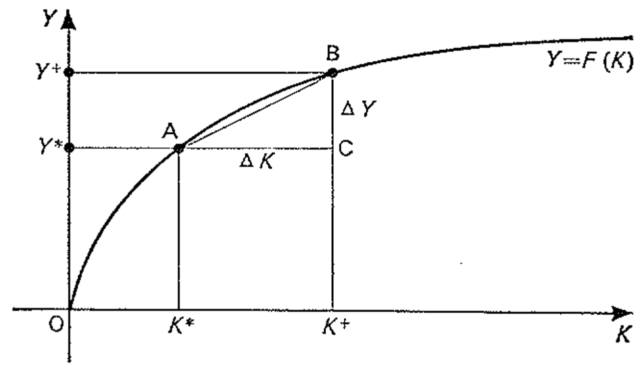

Fig. 2.5

output, rather than output per labourer, is measured on the vertical axis, and the aggregate capital stock, rather than capital per labourer, on the horizontal axis. If the amount of capital in the economy is originally K^* then, given the production function, Y^* of output is produced. Consider an increment ΔK in the capital stock—increasing K to K^+. This generates an increment ΔY in output—increasing Y to Y^+. Now $\Delta Y/\Delta K$, the increment in output divided by the increment in capital, is a rough measure of the marginal product of capital and is measured by the *slope* of the line joining the points AB on the production function (see Lipsey (164) Ch. 2 Appendix). Consider what would happen to this measure if the increment in capital were made smaller and smaller. The marginal product would continue to be measured by the slope of the line AB but the point B would be getting closer and closer to A. If the increment in K were made *infinitesimally* small then it should be clear that the marginal product of capital at the point A will be *measured by the slope of the production function at that point*[19]. Now, the slope of a curve at any point is equal to the slope of the tangent to the curve at that point. Consequently, we can summarize as follows.

In Fig. 2.5, the slope of the tangent at the point A *measures the marginal product of capital, $\partial Y/\partial K$, at that point. If the marginal productivity theory is accepted the same tangent measures the rate of profit or rental rate of*

[19] Many readers will recognize the preceding argument as a version of the fundamental idea of differential calculus. See Allen (6) Chs V and VI.

capital that would be generated by competitive conditions given a quantity of capital of K^*.

The slope of the curve $Y = F(K)$ is continually declining in Fig. 2.5—reflecting the assumption of diminishing marginal productivity.

Consider Fig. 2.6. Given a capital–labour ratio of k^*, an output per labourer of y^* is produced. Now, by analogy with the argument associated with Fig. 2.5, it is not difficult to accept that the slope of the tangent at the

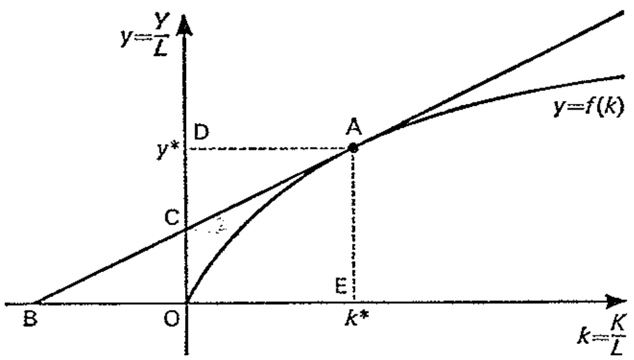

Fig. 2.6

point A measures the marginal product of capital at that point and, if the marginal productivity theory of distribution is accepted, this slope will equal the rate of profit, r. Now, the slope of the line CA is given by CD/DA: i.e.

$$\text{slope of tangent CA} = r = \frac{\text{CD}}{\text{DA}}$$

But DA equals OE which is the capital–labour ratio, k^*, associated with the tangent CA. Consequently,

$$r = \frac{\text{CD}}{\text{OE}} = \frac{\text{CD}}{k^*}$$

or,

$$\text{CD} = rk^* \tag{2.4.7}$$

Now, rk^* is the rate of profit multiplied by the amount of capital per labourer. It is clear therefore that the distance CD measures the amount of *profits per labourer*. Since OD measures the total amount of output per labourer and constant returns to scale, implying, by Euler's theorem, product exhaustion, are assumed, *wages per labourer*, or *the wage rate*, are given by

$$\begin{aligned}\text{Wages per Worker} &= \text{OD} - \text{CD}\\ &= \underline{\text{OC}}\end{aligned} \tag{2.4.8}$$

These results are frequently used in this book and it is essential that the reader should be certain of their derivation. The slope of the intensive production function at a point is *often denoted* as $f'(k)$. Thus, given a capital–labour ratio of k^*, the rate of profit, r, will often be written as

$$r = f'(k^*) \tag{2.4.9}$$

Profits per worker will therefore equal $k^*f'(k^*)$ and wages per worker, w, will, in this notation, be given by

$$\begin{aligned} w &= \text{Output per worker} - \text{profits per worker}\\ &= f(k^*) \qquad\qquad\quad - k^*f'(k^*)\end{aligned} \tag{2.4.10}$$

Finally, it is worth noticing that the slope of the tangent CA can be written in an alternative way

$$\text{Slope of CA} = r = \frac{\text{OC}}{\text{OB}}$$

We have demonstrated that OC measures the wage rate, w. Hence

$$r = \frac{w}{\text{OB}}$$

or,

$$\text{OB} = \frac{w}{r} \tag{2.4.11}$$

Thus, we can summarize the results associated with the use of the per-worker continuous production function in Fig. 2.6 and the assumption of the marginal productivity theory as:

(1) *Profits per worker are measured by the distance* CD.
(2) *Wages per worker are measured by the distance* OC.
(3) *The ratio of wages per worker, w, to the rate of profit, r, is measured by the distance* OB.

The reader would be well advised to experiment with this diagram to confirm his understanding. An investigation of the effect upon w and r of increasing the capital–labour ratio might prove particularly fruitful[20].

(v) THE ELASTICITY OF SUBSTITUTION

This useful concept, usually attributed to J. R. (now Sir John) Hicks ((107) p. 289) can be defined as the proportionate rate of change of the ratio of

[20] A formal mathematical derivation of the results (2.4.9) and (2.4.10) is straightforward. Given $Y = Lf(k)$, the partial derivatives of Y with respect to K and L are obtained and these are the respective marginal products. See Allen (7) pp. 45–6.

capital to labour with respect to a change in the ratio of the prices of capital and labour. In this form the definition sounds rather confusing but the concept is sufficiently widely used to merit an heuristic treatment.

We have already demonstrated that the ratio, w/r, of wages per labourer (the wage rate) to the rate of profit is measured by OB in Fig. 2.6. Consider the effect upon the capital–labour ratio of changing the distance OB while retaining the line AB tangential to the $f(k)$ curve. A *higher* capital–labour ratio is associated with an *increase* in the distance OB, i.e. a *higher* k is associated with a higher value of w/r. Similarly, a *lower* capital–labour ratio is associated with a *decrease* in the distance OB—i.e. a lower k is associated with a lower value of w/r. It is clear that the capital–labour ratio is a function of the wage/rate of profit ratio:

$$k = F\left(\frac{w}{r}\right) \tag{2.4.12}$$

or, inverting $\frac{w}{r}$,

$$k = F\left(\frac{r}{w}\right) \tag{2.4.13}$$

Our previous discussion has demonstrated that as r/w *increases* (and w/r decreases) the capital–labour ratio, k, *decreases*. This relationship is graphed in Fig. 2.7.

The elasticity of substitution, usually denoted σ, is now simply defined as the *elasticity of the curve* AA—by analogy with the conventional idea

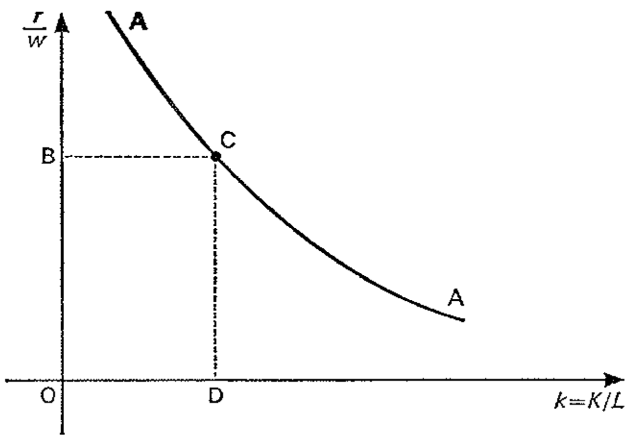

Fig. 2.7

of the elasticity of a demand curve. If we write $\bar{p} = r/w$ then the elasticity of substitution can be defined as

$$\sigma = \frac{\text{Proportionate change in the capital–labour ratio}}{\text{Proportionate change in the factor–price ratio}: \bar{p} = \frac{r}{w}},$$

i.e.

$$\sigma = \frac{\Delta k}{k} \div \frac{\Delta \bar{p}}{\bar{p}} \text{ where } \Delta \text{ means 'a small increment'}$$

or

$$\sigma = \frac{\bar{p}}{k} \cdot \frac{\Delta k}{\Delta \bar{p}} \tag{2.4.14}$$

Because the curve AA is downward sloping, the definition of the elasticity of substitution in (2.4.14) is inherently negative. Following convention the sign is usually reversed:

$$\sigma = -\frac{\bar{p}}{k} \cdot \frac{\Delta k}{\Delta \bar{p}} \tag{2.4.15}$$

to make σ positive[21].

Thus, the elasticity of substitution measures the *responsiveness* of the capital–labour ratio to the prices of capital and labour. If the elasticity of substitution is equal to zero then it is clear that the capital–labour ratio is totally unresponsive to *any* change in r/w—i.e. there is no possibility (either technically or because of the preferences of entrepreneurs) of substituting capital for labour. If, on the other hand, the elasticity of substitution equals one then a small fall in the ratio r/w would be associated with an equal *proportionate increase* in the capital–labour ratio.

It is useful to note the relationship between the elasticity of substitution and the ratio of relative shares in national output of capital and labour[22]. The area of any rectangle, say OBCD, inscribed under the curve AA in Fig. 2.7 is equal to $\frac{r}{w} \times \frac{K}{L} = \frac{rK}{wL}$ or the ratio of relative shares. Now, if the elasticity of substitution is equal to one we know that a 1% *increase* in r/w will be associated with a 1% *decrease* in K/L—and consequently the ratio of relative shares will *remain constant*. If the whole curve AA had an elasticity of one (in which case it would be a rectangular hyperbola—see Lipsey (164) p. 102) then the ratio of relative shares would be the same whatever

[21] Mathematically, σ is simply $-(d \log k)/d \log \bar{p}$. See Allen (7) p. 48, or Brown (31) Ch. 3.

[22] This relationship is formally analogous to the well-known relationship between price elasticity of demand and total household expenditure. See Lipsey (164) Ch. 10.

value of r/w ruled in the economy. If $\sigma > 1$, a 1% increase in r/w will be associated with a decrease in K/L of more than 1% and the ratio rK/wL will therefore fall—the share of capital declining relative to that of labour. Similarly, if $\sigma < 1$, a 1% increase in r/w will be associated with a decrease in K/L of less than 1% and the ratio rK/wL will rise.

(vi) SOME SPECIAL FORMS OF THE CONTINUOUS PRODUCTION FUNCTION
Special forms of the general aggregate production function have become widely used—particularly in empirical studies. We note, for future reference:

The Cobb–Douglas Production Function
This celebrated form of production function is generally attributed to the work of C. W. Cobb and P. H. Douglas in the 1920s (46) although it can be found in the writings of Knut Wicksell (*Finanztheorische Untersuchungen*, 1896, p. 53) and P. H. Wicksteed (278)[23]. In its general form it is written as

$$Y = K^{\alpha}L^{\beta} \tag{2.4.16}$$

If $\alpha + \beta = 1$ then this production function exhibits constant returns to scale and can be written as

$$Y = K^{\alpha}L^{1-\alpha} \tag{2.4.17}$$

In this case, it can be shown (see Allen (7) p. 49 or Brown (31) Ch. 3) that the elasticity of substitution is equal to unity.

The C.E.S. Production Function
This form of the continuous aggregate production function is generally attributed to Arrow, Chenery, Minhas and Solow (9) but was independently derived by Brown and de Cani (30). In its simplest form, the C.E.S. production function is written:

$$Y = [AK^{-\beta} + BL^{-\beta}]^{-1/\beta} \tag{2.4.18}$$

where A and B are constants. This function exhibits the property of possessing a *Constant Elasticity of Substitution*[24] which is given by

$$\sigma = \frac{1}{1 + \beta}$$

[23] For an interesting account of its origin and early hostile reception, see Douglas's essay in Brown (32) pp. 15–22.

[24] See Allen (7) pp. 52–5 or Brown (31) Ch. 3. Notice that both the Cobb–Douglas and Fixed-Coefficient forms of the aggregate production function imply constant elasticities of substitution—but in the Cobb–Douglas case $\sigma = 1$ whereas in the fixed-coefficients case $\sigma = 0$. The C.E.S. form allows σ to take on any value.

It is sometimes called a *homohypallagic production function* from the Greek meaning 'same substitution'. This form of the continuous production function is primarily used in empirical work.

2.5 Rates of Growth

It is obvious that the concept of the *rate of growth* of a variable will be central to a textbook on theories of economic growth. There are a variety of possible ways of defining a rate of growth. The simplest is well known. Consider a labour force of size L. If the labour force grows (as a result of an increase in population, an increase in the participation ratio, or of immigration) by an absolute amount ΔL then we would conventionally define the *proportionate rate of growth of the labour force* as $\Delta L/L$. Thus, the simplest way of defining the rate of growth of a variable is as the ratio of the increment in the variable to the original level. In this sense, $\Delta Y/Y$ would be the proportionate rate of growth of national income and $\Delta K/K$ would be the proportionate rate of growth of the aggregate stock of capital. Although this notation is frequently used in this book, a more precise notation is often more convenient. Notice that our simple definition *does not specify the time period* over which the increment in the relevant variable is measured—i.e. if $\Delta L/L$ is 5%, we do not know whether this refers to a 5% growth per decade, per year, per month or per day. Consider Fig. 2.8 which

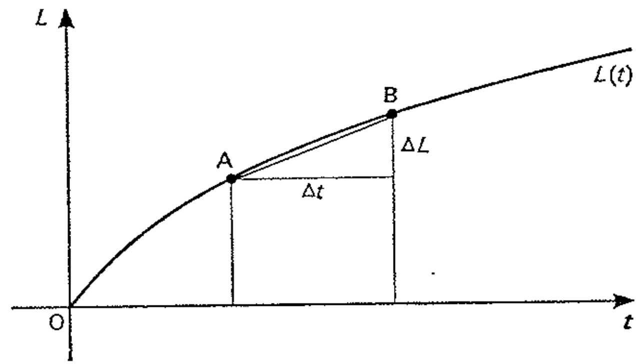

Fig. 2.8

illustrates the growth of the labour force, L, of an economy as time, t, proceeds. If the increment, Δt, in time is a year then ΔL measures the increment in the labour force during the year and $\Delta L/L$ is the proportionate rate of growth of the labour force in the year. We could make our definition of the rate of growth more precise as defining it as: $\frac{\Delta L}{\Delta t} \cdot \frac{1}{L}$ i.e. the *increment per time period* divided by the original level of the labour force. Now, $\Delta L/\Delta t$ is measured by the *slope* of the line AB in Fig. 2.8. Consider what would happen if the increment, Δt, in time was made smaller and smaller

—i.e. corresponding to the increment in the labour force per half-year, per quarter, per month or per day. It is clear that for *infinitesimally small* increments in time we could define the *instantaneous change* in L as the slope of the curve $L(t)$ at the point A. This procedure would imply that we would define the rate of growth of the labour force as $\dot{L}/L$ where the dot over the variable L means *that we are referring to the instantaneous rate of change of the labour force with respect to an infinitesimal increase in time, t.* Although it is important that the reader appreciates the preceding argument, all that is really necessary for understanding this book is the recognition that the following symbols are the precise definitions of the instantaneous proportionate rates of growth of the corresponding variables

$$\frac{\dot{Y}}{Y} = \text{proportionate rate of growth of national income}$$

$$\frac{\dot{K}}{K} = \text{proportionate rate of growth of the capital stock}$$

and, in general,

$$\boxed{\frac{\dot{X}}{X} = \text{rate of growth of the variable } X}$$

This definition of the rate of growth of a variable, which the mathematically trained will recognize as a straightforward application of a concept of differential calculus, recurs throughout the book and the reader will be frequently reminded that a dot over a variable simply signifies the rate of change of that variable with respect to time.

We have defined a concept of the instantaneous proportionate rate of growth. It is sometimes useful to use a mysterious mathematical number called e in this context. Imagine a population of 100 people[25]. They grow at a rate of 6% per year such that at the end of the first year there are $106 = 100\,(1{\cdot}06)$ people. At the end of the second year there are $106 + \frac{6}{100} \times 106$ people, i.e. $100(1{\cdot}06)^2$. In general if the population is growing at 6% per year, then the population after t years will be given by

$$P(t) = 100(1{\cdot}06)^t \qquad (2.5.1)$$

or

$$P(t) = 100(1 + 0{\cdot}06)^t$$

[25] The following discussion relies heavily on Allen (6) pp. 228–9. It can be omitted without any substantial loss of continuity.

which is, of course, the simple compound interest formula used in Section 2.2(b). Thus, in general, an original population P_0 growing at $100n\%$ for t years will amount to

$$P(t) = P_0(1 + n)^t \qquad (2.5.2)$$

Consider now the case where the population is growing at 3% *per half year*. The population at the end of successive *half-years* will be

$$P_0(1{\cdot}03), P_0(1{\cdot}03)^2, P_0(1{\cdot}03)^3 \ldots$$

Hence at the end of *t years* the population will be

$$P(t) = P_0(1{\cdot}03)^{2t} \qquad (2.5.3)$$

Now, in principle, the government statisticians could be measuring the population and calculating the rate of growth, quarterly, monthly or even daily. Thus, in general, a population P_0 growing at a rate of $100n\%$ per year with a compounding being calculated x times a year will grow to

$$P(t) = P_0\left(1 + \frac{n}{x}\right)^{xt} \qquad (2.5.4)$$

This point is most simply seen by considering the effect upon one's savings of interest being added daily rather than once a year. Consider the result of investing a £1 for a year at the rate of interest of 100%. If the interest is compounded *yearly* then the £1 will be doubled at the end of the year. If, on the other hand, it is compounded x times a year then it will amount to

$$£\left(1 + \frac{1}{x}\right)^x \qquad (2.5.5)$$

Now, as x, the number of compoundings becomes larger and larger the total return tends to a definite limit. Thus, if x is 10 then the return is 2·594, whereas if x is a 1,000 then the return is 2·717. As x tends to infinity the return tends to a definite number known as e which is approximately 2·71828. The number e can be usefully employed in the context of any growing variable. Consider expression (2.5.4). It can be rewritten[26] as

$$P(t) = P_0\left[\left(1 + \frac{n}{x}\right)^{x/n}\right]^{nt} \qquad (2.5.6)$$

[26] The rules for manipulating indices are conveniently found in Parry-Lewis (193) p. 39.

If we write $\frac{x}{n} = g$ then (2.5.6) becomes

$$P(t) = P_0\left[\left(1 + \frac{1}{g}\right)^g\right]^{nt} \qquad (2.5.7)$$

Now as x tends to infinity so does g (since $x = ng$). But we know that the expression $\left(1 + \frac{1}{g}\right)^g$ tends to e as g tends to infinity. Hence, expression (2.5.6) tends to: $P(t) = P_0e^{nt}$ as x, the number of compoundings, tends to become very large. Thus, whenever the reader encounters an expression such as $L(t) = L_0e^{nt}$ it should be simply interpreted as stating that the labour force is growing from an original level of L_0 at a constant proportionate rate n. This concept of continuous growth at a constant rate is frequently used in simple growth models—although all real world growth rates are, of necessity, calculated in discontinuous terms.

2.6 Steady Growth

The concept of *equilibrium*, which was probably first encountered by the student in a discussion of simple supply and demand curves, has been central to economic theory since at least the middle of the nineteenth century when the gradual introduction of the mathematical methods of differential calculus provided obvious analogies with the ideas of equilibrium—of bodies at rest and the balance of opposing forces—prevalent in the physical and mechanical sciences. Economic theory uses the idea of equilibrium as a benchmark for the study of an economic process or system. In general terms, it can be taken to refer to that configuration of the economy from which there is no tendency to change. An equilibrium is *stable* if any displacement from the original position produces forces which tend to move the system back to its original position.

Special concepts of equilibrium have been developed for use in the context of a growing economy. These are:

(i) Steady-state Growth

A model economy is said to be experiencing steady growth (is in a steady-state) if all variables are growing at a *constant* proportional rate or not growing at all (i.e. growing at a zero rate)[27].

[27] The definition could, of course, apply to a contracting economy. Allen (7) p. 174 defines steady-state growth in an equivalent way to what we have called balanced growth (see p. 41). Usage of these phrases differs but Hahn and Matthews (85) pp. 3–4 is a useful authoritative reference.

Thus, in a state of steady growth, the capital stock will grow at a *constant* proportional rate—say 5% per annum and an increasing or decreasing rate of growth violates our definition of steady growth. It is important to realize that the idea of steady growth is a convenient method whereby the economist hopes to analyse *some* of the issues associated with a growing economy. It is *not* a name for some phenomenon of observed reality although Solow has argued

> 'Real economies are not in steady states; they are not in any state that can be described in a word. But they do not appear to be very far from, or to be rushing systematically away from, steady-state conditions. So the steady-state may be a fair first approximation' (Solow in Burmeister and Dobell (34) pp. vii and viii).

He is careful to add 'that this is a temporary excuse not a permanent licence'.

Many models of economic growth are concerned to identify whether a state of steady growth can possibly be achieved—an investigation which we will refer to as the *Existence* problem. Another question that is frequently investigated is whether there are forces in the economy which tend to drive the system towards the steady state. We will refer to this as the STABILITY PROBLEM[28].

(ii) Balanced Growth

We will describe an economy as experiencing balanced growth if all the variables are growing at the *same* constant rate or not growing at all. This definition will require some slight modification as the book proceeds but the general idea should be clear. In a state of balanced growth the main aggregative variables remain in the same proportion to each other. In many of the models that we will be discussing a steady-state path will also be a balanced growth path and we will be able to use the terms interchangeably when there is no risk of confusion. As with the steady state we will be interested in the Existence and Stability of balanced growth paths[29]. The concentration of most modern writers on economic growth on the properties of paths of *balanced growth* has been subject to increasingly severe criticism—and many of the arguments are discussed as this book proceeds. Professor Mrs Robinson has referred to a path of growth involving

[28] There are different aspects to the stability problem. See Hahn and Matthews (85) pp. 3 and 4.

[29] It should be noted that this definition of balanced growth is *different* from that used in other areas of economics—notably development economics. It is equivalent to Hicks' [(108) p. 133] definition of 'growth equilibrium' in which 'all elements in the economy are growing at the same (constant) rate—so that although there is an absolute expansion every element remains in the same proportion to every other'.

balanced growth and full employment of *labour* as a 'Golden Age' 'thus indicating that it represents a mythical state of affairs not likely to obtain in any actual economy' (Robinson (209) p. 99).

The meaning of the above ideas will become clearer when they are put to use in the context of specific models of economic growth—an activity which we are, at last, in a position to begin.

chapter 3
The Harrod–Domar Model of Economic Growth

3.1 Introduction

Contemporary interest in *modern* theories of economic growth can be conveniently dated from the publication of Harrod's seminal paper (99) followed shortly by Domar's similar, but independently derived, contributions (57) (58)[1]. Having examined some of the more important concepts and methods of growth theory in Chapter 2 we are, at last, in a position to begin our study of the theories themselves and both tradition and the need for simplicity dictate that the most convenient starting point is the approach which has come to be known as the Harrod–Domar theory of growth.

This chapter will consist of an exposition of the *simplest* version of Harrod's theory. Two important points should be noted.

(a) The simple approach to Harrod's theory outlined here is not, and could not be, an exact representation of Harrod's thought. There are a host of possible formalizations which claim to capture the essence and spirit of Harrod's contribution[2] and, as might be expected, his own views have been modified and extended in the years since the publication of his justly celebrated *Towards a Dynamic Economics*. It seems clear that Harrod's theorizing was far more subtle and sophisticated than would be concluded from a reading of some of the simplest stylizations of the exegetical literature and Harrod himself has expressed irritation at what he considers to be *misrepresentation of some of his ideas*[3]. Some aspects of this problem of interpretation will be discussed later in the chapter while a discussion of the complete Harrod model, incorporating the effects of technical progress, is postponed until Section 7.4. Our purpose here is to try and isolate some of the *central* issues associated with the Harrod model and attempts to do this while discarding, or lessening the emphasis on, particular aspects can be partly justified by reference to Harrod himself:

[1] Harrod's 1934 paper (96) and his 1936 book (97) both include many of the central ideas of his later theorizing. Lundberg's fascinating book (168), which predated Harrod's explicit attempts at dynamic theorizing, contains many fascinating insights into the process of growth in a market economy.

[2] See, for example, Ackley (2), Alexander (5), Baumol (17) and Hicks (106).

[3] See, for example, his reply to Mrs Robinson in (103).

> 'The significance (of Harrod's analysis) should not be judged solely by reference to the validity or convenience of the particular equations set forth. It involves something wider: a method of thinking, a way of approach to certain problems. It is necessary to think dynamically.' (Harrod (99) p. 15).

(b) The greater emphasis given to Harrod's approach should not be taken as a reflection on the relative importance of Domar's work—and any serious student of the theory of growth should, at a minimum, read his original contributions (57) (58). However, Domar's purpose was somewhat more limited than that of Harrod and, in particular, the question of the stability of a growing capitalist economy was not a central feature of the analysis. Thus, the best policy would appear to be a concentration on Harrod's theory, with suitable digressions to discuss Domar's work. The similarities and differences between the two theories are discussed in Section 3.7.

Harrod's approach to growth, as outlined in (99) and elaborated in his lectures at London University (100) is particularly Keynesian both in the spirit of the conception and in the details of execution[4]. Building on the concepts and methods of Keynesian short-run macroeconomics, it concentrates on the necessary conditions for equilibrium between aggregate saving and investment in a dynamic economy. However, at the time that Harrod wrote, the concept of the *rate of growth* of the major components of the macroeconomy or, for that matter, of any economic variables, lived a curious half-life of acknowledgement in footnotes but exclusion from the main body of economic abalysis[5]. As Domar commented in 1952:

> 'In economic theory growth has occupied an odd place: always seen around but seldom invited in. It has either been taken for granted or treated as an afterthought.' (Domar (59) p. 16).

Given the post-war concern with high rates of economic growth and, latterly, with possible 'limits to growth', it is difficult to comprehend the relatively short public life of a theoretical interest in the causes and mechanisms of economic growth.

[4] Harrod was one of the economists who commented in detail on prepublication drafts of *The General Theory*—see the list of letters on pages 517–8 of (143). Harrod's original 'essay in dynamic theory' (99) was subjected to sustained criticism by Keynes who was, at that time, the editor of the *Economic Journal*. See pp. 321–50 of (143).

[5] Although earlier writers had acknowledged, in general terms, the need for a dynamic economics. Thus, for example, J. B. Clark, one of the originators of neoclassical economics (see 4.1), emphasized that 'A theory of disturbance and variation' would be 'included in the science of economic dynamics but the most important thing that is included in it is a theory of progress'. See (42) pp. 31 and 33.

3.2 The Harrod Model

Harrod's 'dynamic theory' can be alternatively viewed as simple or subtle—some would say over-simple or over-subtle. A central aim was the construction of the 'fundamental dynamic principle' whose derivation is discussed below. This principle was attractive to Harrod partly because of its 'extreme simplicity' (Harrod (100) p. 80) and partly because he knew 'of no alternative formulation, in the world of modern economic theory, of any dynamic principle of comparable generality' (Harrod (100) p. 80). On the other hand, Harrod, like many economists of his era, combined the quest for theoretical simplicity with a taste for descriptive realism and immediate policy application. Thus, his 1947 lectures included not only abstract macroeconomic theorizing but also detailed discussions of the motives of real microeconomic agents and skilful application of his new theories to the pressing economic problems of the United Kingdom in the immediate post-war period. Harrod's early work on economic growth can delight or infuriate—depending upon the temperament and training of the reader. It can be seen as a brilliant tour de force—exemplifying the best features of the tradition of political economy, or as a rather vague and woolly piece of economic theorizing which, for all its insights, exhibits a lack of rigour which weakens its conclusions. But, whichever interpretation is preferred, Harrod's work was, and is, impossible to ignore.

Harrod's *formal* analysis is conducted within a highly aggregative framework although the implied aggregation procedure is at no stage made explicit. This was, of course, the usual approach following Keynes' rehabilitation of aggregative economics although, as we have already noted, (see 2.2) Keynes himself was invariably careful when using aggregative concepts. Harrod's macro variables are 'heroic' aggregates (see Section 2.2(d)) and it would be tempting to interpret his model as referring to an economy in which only one good is produced—thereby completely avoiding the aggregation difficulties discussed in Chapter 2. However, resorting to the 'one-good parable' would undoubtedly misrepresent much of the flavour, and some of the central results, of the Harrod model. A possible way out of the aggregation difficulty is to assume that relative prices are constant[6]—in which case the principal macro variables of the Harrod model can be interpreted as value-based aggregates. In Hicks' words:

> 'When prices are constant, quantities of goods and services can be added by adding their money values; money values become volume indexes' (Hicks (108) p. 78).

Such an assumption is clearly unsatisfactory in the context of a general theory of economic growth but it does give a more precise meaning to the

[6] Hicks has dubbed this approach 'the Fixprice method' and notes that it is 'mixed up with the so-called "Keynesian revolution" '. For an excellent discussion see Ch. VII of Hicks (108).

aggregates in the Harrod model. The problems of aggregation are returned to in Section 6.2.; for the moment it is sufficient that the student recognize that they exist.

The explicit assumptions of the simplest version of Harrod's model can be quickly outlined.

ASSUMPTION 3.2.1

Savings, S, are assumed to be a simple proportional function of national income, Y. (See 2.3a): $S = sY$ where $s =$ the average and marginal propensity to save. Harrod did not himself assume that s was constant and in his long analysis of aggregate saving (Ch. 2 of (100)) he noted that 'in crucial cases saving as a fraction of income might not be constant' ((100) p. 79). On the other hand, it does no great violence to Harrod's approach to explicitly assume a constant average propensity to save.

ASSUMPTION 3.2.2

The labour force, L, is assumed to grow at a constant exogenous rate n: $\dot{L}/L = n$. (See Sections 2.2(c) and 2.5). The assumption of exogeneity implies that the rate of growth of the labour force is completely uninfluenced by other components of the economic system. This assumption diverges sharply from the 'classical' tradition and all Malthusian notions, of the kind studied in elementary courses in economic history and economic development, are explicitly rejected.

ASSUMPTION 3.2.3

There is no technical progress[7] and the capital stock, K, does not depreciate. Neither of these assumptions is necessary for the development of a Harrod-type model of economic growth. They are employed here merely for simplicity.

ASSUMPTION 3.2.4

The amounts of capital, K, and labour, L, required to produce any given flow of output, Y, are *uniquely* given. Thus, the production function implied by the Harrod approach is of the fixed proportions variety discussed in Section 2.4(a):

$$Y = \min \left[\frac{K}{v}, \frac{L}{u}\right] \tag{3.2.1}$$

It is necessary to be particularly careful in interpreting the fixed capital–output and labour–output ratios implied by the Harrod model and the

[7] The Harrod model including the effect of technical progress is discussed in Section 7.4.

matter is extensively discussed in Section 3.5. For the moment we can confine ourselves to an examination of the implications of these assumed constancies.

(a) *Labour* With u defined as the *constant* ratio of labour requirements to total output, it is clear that the production of any given flow of output requires L/u units of labour. Put another way, if all labour is fully employed then the *maximum* flow of output, *whatever* the size of the stock of capital, (see 2.4(a)) is L/u. If, however, the labour force is growing (as is assumed in (3.2.2) above) then the maximum available flow of output *can* grow but a little reflection will confirm that, given the assumption of a constant labour/output ratio, the rate of growth of income or output, $\dot{Y}/Y$, cannot permanently exceed the rate of growth of the labour force which is, by assumption, a constant n. If, at the beginning of time, all labour is fully employed, this assumption implies that, in the absence of technical progress, the maximum rate of growth of national income and output is given by the exogenously determined rate of growth of the labour force.

(b) *Capital* The capital–output ratio, v, in the form implied by equation (3.2.1), is simply the ratio of the capital stock to the flow of output or income—i.e. K/Y[8]. Harrod, however, was primarily concerned with the *marginal* capital–output ratio—i.e. the increment in capital[9] stock associated with an *increment* in output.

If we write

$$K = vY \tag{3.2.2}$$

then, for small increments ΔK and ΔY, it follows that

$$\Delta K = v\Delta Y \tag{3.2.3}$$

or, using the rate of change notation introduced in Section 2.5,

$$\dot{K} = v\dot{Y} \tag{3.2.4}$$

We assume that the average capital–output ratio, K/Y, equals the marginal capital–output ratio, $\Delta K/\Delta Y$, although Harrod did not make this assumption explicitly.

It is crucial to notice that *two* different conceptions of the marginal capital–output ratio can be distinguished.

[8] Harrod used the symbol C instead of v. Most of the more recent literature uses the symbol v for the capital–output ratio and we follow this practice.

[9] Harrod emphasized that v referred to 'the addition to capital, but need not consist exclusively or even mostly of capital goods. It is merely the accretion during the period of all goods'. It is simpler and more convenient if this point is ignored and v is taken to refer exclusively to the addition to the capital stock.

$I_t = v(\Delta Y_t)$

(i) *The actual* increment in the capital stock in any period divided by the *actual* increment in output. Thus, at the end of a year, v could be interpreted as the measured increase in the capital stock during the year divided by the measured increase in income or output. We will refer to this interpretation as Definition (i).

(ii) The increment in the capital stock associated with an increment in output that is *required* by entrepreneurs if, at the end of the period, they are to be satisfied that they have invested the correct amount—i.e. if the new capital stock is to equal the amount that they consider appropriate for the new level of output and income. We will refer to this interpretation as Definition (ii) and use the symbol v_r to distinguish this conception from the first[10]. The significance of the two conceptions of the marginal capital–output ratio will become clear at the end of this section.

Assuming, as we have, (see Assumption 3.2.3) that the capital stock does not depreciate, then $\dot{K}$, the rate of change of the capital stock, will, if positive, equal the flow of aggregate investment, I, and equation (3.2.4) can be rewritten as

$$I = v\dot{Y} \tag{3.2.5}$$

which, in that it relates aggregate investment to the rate of change of nation income or output, can be seen to be a simple form of the 'accelerator' mechanism discussed in Section 2.3(b)[11].

Given the above assumptions it is a relatively easy matter to derive the central conclusions of Harrod's analysis of economic growth. The reader will recall the familiar equilibrium condition of elementary macroeconomics

$$I = S \tag{3.2.6}$$

i.e. aggregate planned investment must equal aggregate planned saving. Given the proportional savings function (Assumption 3.2.1) and the accelerator relation of equation (3.2.5), the equilibrium condition of equation (3.2.6) can be rewritten as

$$v\dot{Y} = sY$$

or,

$$\boxed{g = \frac{\dot{Y}}{Y} = \frac{s}{v}} \tag{3.2.7}$$

[10] Harrod used the symbols C_p (1939 paper) and C (1948 book) to refer to Definition (i) and the symbols C (1939 paper) and C_r (1948 book) to refer to Definition (ii).
[11] Harrod's preference was to call this mechanism 'the relation'. See footnote 2 of (99).

● From equation (3.2.7)

$$\ln Y(t) = \int \frac{s}{v}\,dt \qquad (\ln = \text{natural logarithm})$$

$$= \frac{s}{v}t + Z \qquad (Z = \text{constant of integration})$$

Hence,

$$Y(t) = \exp\left[\frac{s}{v}(t) + Z\right]$$

$$= \exp Z \exp \frac{s}{v}t = Y(0) \exp \frac{s}{v}t = Y(0)\, e^{gt}$$

where $Y(0)$ denotes the given initial level of income, and, for convenience, we have used the symbol 'exp' instead of 'e'. Thus, for example, $e^{gt} = \exp(gt)$. ●

Now, $\dot{Y}/Y$ is the rate of growth of national income or output (see 2.5) and equation (3.2.7), which Harrod called the 'fundamental' equation, shows that it must equal the ratio of the propensity to save, s, to the capital–output ratio, v, if equilibrium between aggregate saving and aggregate investment is to be maintained as time goes on. Moreover, since both s and v are, by assumption, constant, the rate of growth of national income must be *constant*. However, as we will see below, some care is required in the interpretation of this equation.

The growth rate of the capital stock is easily derived. Since we are assuming the absence of depreciation, I can be replaced by $\dot{K}$ in the equilibrium condition (3.2.6):

$$\dot{K} = S$$

or, utilizing the proportional savings function,

$$\dot{K} = sY$$

Replacing Y by K/v we obtain

$$\dot{K} = \frac{s}{v}K$$

or,

$$\boxed{\frac{\dot{K}}{K} = \frac{s}{v}} \tag{3.2.8}$$

● Using exactly the same procedure as in the last mathematical digression of this section, it can be shown that:

$$K(t) = K(0) \exp \frac{s}{v} t$$

where $K(0)$ denotes the given initial capital stock. ●

We have therefore demonstrated that both national income, Y, and the capital stock, K, must both grow at the *same constant rate*, s/v,—a situation that corresponds to our definition of steady-state growth[12]. (See Section 2.6.)

The 'fundamental' equation (3.2.7) can be interpreted in two different ways—depending upon which conception of the marginal capital–output ratio (see Definitions (i) and (ii) above) is employed.

(a) The Fundamental Equation as a Truism

Consider equation (3.2.7):

$$\frac{\dot{Y}}{Y} = \frac{s}{v}$$

or,

$$\frac{\dot{Y}}{Y} v = s \tag{3.2.9}$$

If we interpret the marginal capital–output ratio, v, in terms of Definition (i)—as the ratio of the *actual* rate of change of the capital stock (i.e. actual investment) to the *actual* rate of change of national income and output (i.e. $v = \dot{K}/\dot{Y} = I/\dot{Y}$ (then equation (3.2.9) can be written as

$$\frac{\dot{Y}}{Y} \cdot \frac{I}{\dot{Y}} = s = \frac{S}{Y}$$

and, by cancelling the $\dot{Y}$'s and multiplying both sides of the equation by Y, this reduces to the familiar accounting identity that investment, I, *must* equal savings, S, *ex-post*[13]. If the marginal capital–output ratio, v, is given this interpretation then the fundamental equation is a 'truism'—it is 'necessarily true' and 'follows from the definition of the terms' (Harrod (100) p. 80). To labour the point, if v is defined as in Definition (i) then the rate of growth of national income *must* equal s/v. Using the symbol G_A for the *actual* rate of growth of national income over any period of time, the fundamental equation, viewed as a truism, can be written as

$$G_A \equiv s/v \tag{3.2.10}$$

[12] Simple manipulation of the variables will confirm that investment grows at the same rate, s/v.

[13] See the discussion of the distinction between planned and actual quantities in Section 2.2.

where the identity sign, $\equiv$, reminds us that the marginal capital–output ratio is defined in such a way as to make the statement definitionally true.

(b) The Fundamental Equation Defining an Equilibrium Growth Path

The 'fundamental equation' can, however, be given theoretical content if the marginal capital–output ratio is interpreted in the second of the two ways discussed above. (Definition (ii))—i.e. as expressing the entrepreneurs' *requirements* for additions to the capital stock given the growth of income and output. Using the symbol v_r introduced above (p. 48) we can write

$$\frac{\dot{Y}}{Y} = \frac{s}{v_r} \tag{3.2.11}$$

Equation (3.2.11) is no longer a truism. It expresses the rate of growth of output, which we denote G_w, which will satisfy entrepreneurs that they are investing the correct amount. Equations (3.2.10) and (3.2.11) imply that

$$G_A v = s = G_w v_r$$

Now, if the actual rate of growth that occurs, G_A, equals the necessary growth rate, G_w then it is clear that, v the actual marginal capital–output ratio, must equal v_r, the required marginal capital–output ratio. Put another way, if national income and output happens to grow at the rate G_w then the *actual* increase in the capital stock associated with the growth of income will equal the increase that entrepreneurs *require* if they are to be satisfied that the level of the capital stock is exactly appropriate for the production of the current level of national output. Harrod called the rate of growth G_w the '*warranted*' rate and defined it as 'that overall rate of advance which, if executed, will leave entrepreneurs in a state of mind in which they will be prepared to carry on a similar advance' (Harrod (100) p. 82). It is easy to see that, if output actually grows at the warranted rate, then the actual capital stock will conform to the desired capital stock—and a wide range of assumptions on the behavioural responses of entrepreneurs would imply that, this being so, they would be prepared to carry on the same rate of growth in the future.

In elementary microeconomics we refer to a situation in which entrepreneurs have no incentive to change the ruling price and output of their product as an *equilibrium* configuration. It is clear that Harrod's conception of a 'warranted' advance is a particular notion of *equilibrium growth* in that, if the economy happens to grow at the warranted rate, there is no obvious incentive for entrepreneurs to try and increase or decrease the overall rate of growth of output. The use of the word 'equilibrium' appeared inappropriate to Harrod because of the particular properties of the warranted rate which are discussed in Section 3.4.

3.3 The First Harrod Problem

Thus far, we have used some simple algebraic manipulation to demonstrate that macroeconomic equilibrium in a Harrod-type model of the economy implies a constant rate of growth of output and capital at the warranted rate $G_w = s/v_r$. There is not, of course, any particular reason why we should expect that the economy will actually grow at the warranted rate—the actual rate of growth being the outcome of the expectations, decisions and mistakes of a host of different decision-makers. On the other hand, we have seen that if G_A does not equal G_w then it must be the case that the actual capital stock will not equal the desired capital stock that entrepreneurs consider appropriate.

The level of employment has not yet entered our scheme although we might expect this to be central to what we have described as a 'Keynesian' model. In setting out the assumptions of the model we noted that the actual rate of growth of output could not *permanently* exceed the rate of growth of the labour force because *of the assumed constancy of the labour–output ratio*[14].

Thus, $$G_A \leqslant \frac{\dot{L}}{L} = n$$

Now, if the economy is originally in a situation of full employment, full employment *through time* would imply that the actual rate of growth, G_A, would *equal n*. But we have already seen that, for *equilibrium steady-state* growth, G_A must equal G_W. It is therefore clear that equilibrium steady growth with full employment *necessitates* that

$$G_A = G_W = n \qquad G = \frac{\dot{L}}{L} = \frac{\dot{K}}{K}$$

or, $$G_A = \frac{s}{v_r} = n \tag{3.3.1}$$

If equation (3.3.1) is satisfied then the economy will grow at a constant proportional rate of $s/v_r = s/v = n$ a situation which Mrs Robinson has described as 'The Golden Age'—'thus indicating that it represents a mythical state of affairs not likely to obtain in any actual economy' (Robinson (209) pp. 99–100). It is therefore clear that the Harrod model includes the *possibility* of equilibrium steady growth at full employment.

However, there is clearly no reason to believe that s/v_r *will* equal s/v or n. s, v and n are all independently determined. Only a 'happy accident' ((85) p. 7) will generate steady-state growth at full employment in the Harrod model. The propensity to save, s, is determined by the preferences of firms

[14] If the economy were emerging from a depression involving a substantial amount of unemployment of resources then it would, of course, be possible for the actual growth rate to *temporarily* exceed the rate of growth of the labour force.

and households in the economy. The rate of growth of the labour force, n, is exogenous to the economic system—determined simply by the biologically determined birth and death rates. The capital–output ratio, v, is, on our present interpretation, a reflection of the fixity of the technology. If, by coincidence, the actual rate of growth equalled the warranted rate, which itself equalled the rate of growth of the labour force, then steady growth at full employment would occur. But, *there is no mechanism in the Harrod Model which would ensure the attainment of this Golden Age situation.* Harrod referred to the rate of growth of the labour force (in the absence of technical progress) as the 'Natural' rate of growth[15]. For future reference we can summarize what we will call the First Harrod Problem as follows.

First Harrod Problem

Although steady state growth at full employment is possible *in an Harrod-type model of economic growth, such a 'Golden Age' is highly* improbable *given the independent constituent variables in the necessary equality of the warranted rate of growth, s/v_r, to the natural rate of growth, n.*

This conclusion is thoroughly 'Keynesian' in spirit: there is no reason to believe that full-employment equilibrium growth will be attained. Thus, the 'first Harrod problem' can be interpreted as a dynamic version of the central Keynesian allegation that under-employment equilibrium is possible in a capitalist economy. Much of the literature on the theory of growth in the last twenty years is capable of being interpreted as a sustained attempt to weaken this conclusion.

3.4 The Harrod Stability Problem

The 'First Harrod Problem' is only a first step towards what Harrod regarded as his main theme: 'that sooner or later we shall be faced once more with the problem of stagnation' (Harrod (100) p. v). In a well-known, if frequently misunderstood, argument, Harrod suggested that the warranted rate of growth was *fundamentally unstable* in the sense that divergences of the actual rate of growth, G_A, from the warranted rate, G_w, would not only *not* correct themselves but would produce even larger divergences. Harrod provided what he considered to be 'an extraordinarily simple and notable demonstration of the instability of an advancing system' in the sense that 'around the line of advance which, if adhered to, would alone give satisfaction, centrifugal forces are at work, causing the system to depart farther and farther away from the required line of advance' (Harrod (100) p. 86).

[15] It should be emphasized that Harrod's terminology does not imply that this is the natural rate of growth in the sense of 'normal' or 'usual'. In particular, one should not infer that this is the rate of growth that will be generated by the free operation of market forces.

Harrod's argument *is* simple. We have already noted that equations (3.2.10) and (3.2.11), taken together, imply that

$$G_A v = s = G_w v_r \tag{3.4.1}$$

and therefore that the actual rate of growth, G_A, will equal the warranted rate, G_w, if, and only if, the actual marginal capital–output ratio, v, equals the required capital–output ratio, v_r. It is clear from equation (3.4.1) that if G_A *exceeds* G_w then v_r will exceed v. Conversely, if G_w exceeds G_A then v will exceed v_r. This is the essence of Harrod's instability problem. If the actual rate of growth happens to exceed the warranted rate, entrepreneurs will find that the increase in the capital stock that actually occurs is *less* than the increase that they require given the growth of income and output. Harrod envisages their response to this discrepancy to be an attempt to invest even more which will, of course, force the actual rate of growth even further above the warranted rate and actually *increase* the discrepancy between the actual and desired capital stock.

The difficulty with Harrod's conception of instability is that it is not quite clear what he means[16] and this factor has provided plenty of scope for a host of contradictory interpretations of the stability problem. Some, like Rose (215), have reached conclusions diametrically opposite to those of Harrod, and Hahn and Matthews pointed out that

> 'The instability or otherwise of the system depends on the exact error-adjustment assumptions made. Some formalizations of the model suggest Harrod's main conclusions, while others do not, and yet others conclude that it depends on the exact values taken by the parameters.' (Hahn and Matthews (85) p. 27).

A particularly simple and illuminating formalization of the instability problem has been provided by Sen ((237) pp. 11–13) and his version has the added merit of highlighting the central role in the analysis of the exact expectational pattern ascribed to entrepreneurs.

It is necessary to define some additional symbols:

Y_t^E : represents the flow of output and income that entrepreneurs *expect* in period t.
Y_t : represents the *actual flow* of output and income in period t.
G_t^E : represents the *expected* rate of growth of output and income from period $t-1$ to period t. It is defined as:

$$G_t^E = \frac{Y_t^E - Y_{t-1}}{Y_t^E}$$

[16] See, for example, Jorgenson (119). Harrod's new book (104) does, in fact, clarify some of the issues involved. See Ch. 3.

G_t : represents the *actual* rate of growth of output and income from period $t-1$ to period t. It is defined as[17]:

$$G_t = \frac{Y_t - Y_{t-1}}{Y_t}$$

Through the usual simple multiplier process of elementary macroeconomics, investment at time t determines the actual level of income at time t:

$$Y_t = \frac{1}{s} I_t \tag{3.4.2}$$

where, as in the discussion of Sections 3.2 and 3.3, s represents the propensity to save.

A very simple accelerator mechanism determines the level of investment in period t:

$$I_t = v(Y_t^E - Y_{t-1}) \tag{3.4.3}$$

i.e. investment in period t is a simple proportion of the expected additional output required and the capital–output ratio, v, is seen in its role as an accelerator coefficient. (See Section 2.3.ii.) It is easy to visualize this process in terms of an entrepreneur who determines his capital requirements for the coming year on 1st January. If he expects demand for his product to be greater than the actual demand in the previous year he will, on the assumption that all his machinery was fully employed in the previous year, obtain new machinery to produce the additional required output.

If equation (3.4.3) is substituted into equation (3.4.2) we obtain

$$Y_t = \frac{v}{s}\left(Y_t^E - Y_{t-1}\right)$$

and, dividing both sides of the equation by Y_t^E, we can write

$$\frac{Y_t}{Y_t^E} = \frac{v}{s}\left[\frac{Y_t^E - Y_{t-1}}{Y_t^E}\right] \tag{3.4.4}$$

Now, the expression within the brackets in equation (3.4.4) is, by our definition, the expected rate of growth of income and output. Hence, equation (3.4.4) can be written as

[17] Notice that both the expected and actual growth rates are, for convenience, defined somewhat unconventionally in that the denominator of both expressions is in terms of the expected and actual flows of output in period t rather than $t-1$.

$$\boxed{\frac{Y_t}{Y_t^{\mathrm{E}}} = \frac{v}{s} \cdot G_t^{\mathrm{E}} \qquad (3.4.5)}$$

If entrepreneurs' expectations of the flow of output in period t are perfectly realized then Y_t will equal Y_t^{E} and, as a consequence, $G_t^{\mathrm{E}} \cdot \frac{v}{s}$ will equal 1. It is therefore clear that expectations will be realized if, and only if, the *expected* rate of growth of output happens to equal s/v which, in Harrod's terminology, is the warranted rate of growth. Equation (3.4.5) adds clarity to Harrod's definition of the warranted rate. If entrepreneurs expect a rate of growth of output equal to the warranted rate then their expectations will be fulfilled and many simple psychological mechanisms would suggest that they would be 'prepared to carry on a similar advance'.

We need, however, to consider the effects upon the system if entrepreneurs *do not* expect a rate of growth equal to the warranted rate. Some rather messy substitution [18] generates the following relationship between the actual and expected rates of growth of income and output:

$$\boxed{G_t = 1 - \frac{(1 - G_t^{\mathrm{E}})}{G_t^{\mathrm{E}}} \cdot \frac{s}{v} \qquad (3.4.6)}$$

An examination of this relation demonstrates that the actual rate of growth, G_t, will equal the expected rate of growth if, and only if, the expected rate of growth equals the warranted rate. [19] Moreover, if the expected rate of growth exceeds the warranted rate, then the actual will exceed the expected! We can summarize the deductions from equation (3.4.6) as follows:

$$\left.\begin{array}{l} \text{If } G_t^{\mathrm{E}} > s/v \text{ then } G_t \text{ will be greater than } G_t^{\mathrm{E}} \\ \text{If } G_t^{\mathrm{E}} = s/v \text{ then } G_t \text{ will be equal to } G_t^{\mathrm{E}} \\ \text{If } G_t^{\mathrm{E}} < s/v \text{ then } G_t \text{ will be less than } G_t^{\mathrm{E}} \end{array}\right. \qquad (3.4.7)$$

[18] From the definitions of the actual and expected rates of growth simple manipulation shows that

$$Y_t^{\mathrm{E}} = \frac{Y_{t-1}}{(1 - G_t^{\mathrm{E}})} \text{ and } Y_t = \frac{Y_{t-1}}{(1 - G_t)}$$

Substitution of the above expressions into equation (3.4.5) together with some rearrangement will generate the desired result.

[19] It is easy to check this assertion by inserting appropriate numbers into equation (3.4.6). For example, if the expected growth rate $= s/v = 10\%$ or 0.1 then

$$G_t = 1 - \frac{(1 - 0{\cdot}1)}{0{\cdot}1}0{\cdot}1$$
$$= 0{\cdot}1$$

As Sen comments: 'this is the beginning of Harrod's instability problem' ((237) p. 12). If entrepreneurs coincidentally expect a rate of growth equivalent to the warranted rate then their expectations will be realized. If, however, they expect a rate higher than s/v then the actual rate will exceed their (objectively) optimistic expectations and it seems reasonable to suppose that they might expect an *even higher* rate of growth in the next period. Conversely, if the rate of growth that they expect is less than the warranted rate then the actual rate will be below their expectations and they are likely to feel that they had been over-optimistic and, consequently, revise their expectations downwards in the next period.

It is important to note that, although the seeds of the instability problem are implicit in equation (3.4.6), it is necessary to specify the behavioural responses of the entrepreneurs before one can assert unequivocally that the warranted rate of growth is unstable. In the situation described by equation (3.4.6), a large set of plausible expectations-generating mechanisms will produce a highly unstable system and confirm Harrod's suggestion of the centrifugal forces operating in the vicinity of the warranted rate. If, for example, entrepreneur's expectations are generated as follows:

$$\begin{array}{lll} G_t^{\mathrm{E}} > G_{t-1}^{\mathrm{E}} & \text{if} & G_{t-1} > G_{t-1}^{\mathrm{E}} \\ G_t^{\mathrm{E}} < G_{t-1}^{\mathrm{E}} & \text{if} & G_{t-1} < G_{t-1}^{\mathrm{E}} \end{array} \qquad (3.4.8)$$

then the system is clearly highly unstable—a small deviation of the actual rate from the warranted rate will generate cumulative effects with the actual rate of growth deviating ever further from the steady-state rate of growth s/v. If, for example, s was 20% and v 2 then the warranted rate of growth would be 10%. If entrepreneurs as a whole expect income and output to grow by 10% then it will be necessary for them to invest, and the volume of their investment will be determined by the investment function (equation (3.4.3)) and the accelerator coefficient of 2. Given the flow of investment so generated, the actual level of income will be determined through the multiplier process (equation (3.4.2)) and, from equation (3.4.6), we know that the actual growth in income will, in fact, be exactly 10%. If, on the other hand, entrepreneurs invest on the basis of an expected growth in income of 11% then the actual growth in income will be higher—say 13%. Given the expectational mechanism embodied in (3.4.8), it is clear that in the next period they will invest on the basis of an expected growth rate of somewhat higher than 13% and, once again, the actual growth rate will transpire to be even higher. Thus, deviations from the warranted rate of growth of 10% are not self-correcting and the instability results from the *interaction* of the system described by equations (3.4.2) and (3.4.3) with the expectational mechanism incorporated in (3.4.8).

Harrod's original article included a statement which appeared paradoxical in the extreme but which is easily explained when the ideas of instability discussed above are understood: 'a condition of general overproduction is the consequence of producers in sum producing too little' (Harrod (99) p. 24).

General overproduction occurs when entrepreneurs find that they are unable to sell all the goods that they have produced and consequently 'find themselves in possession of an unwanted volume of stocks or equipment' (Harrod (99) p. 24). Put another way, the *actual* growth in income and the demand for output has been less than the *expected* growth on which they based their output and investment decisions. But, from (3.4.7), we know that this can only occur if $G_t^E < s/v$—i.e. had they expected a higher growth rate of s/v and invested *more* then the overproduction would not have occurred! Hence Harrod's statement.

It should be clear that there is *no* reason why entrepreneurs' expectations should be consistent with the warranted rate of growth. They have no means of knowing s/v and there would be no reason for them to suppose that a consideration of this expression should enter into their decision-making process. Thus, we can outline what we will call the 'Second Harrod Problem'.

Second Harrod Problem

Deviations of the actual rate of growth of a Harrod-type economy from the warranted rate, s/v, far from being self-correcting, are cumulative in effect.

The Second Harrod Problem is frequently referred to as the 'knife-edge' property, a description which graphically captures 'the characteristic and powerful conclusion of the Harrod–Domar line of thought . . . that even for the long run, the economic system is at best balanced on a knife edge of equilibrium growth' (Solow (244) p. 65). In his recent writings Harrod has objected to this phrase complaining that 'Nothing that I have ever written (or said) justifies this description of my view' ((104) p. 32). He dislikes the 'knife-edge nomenclature. . . . because it sounds utterly unrealistic and even a trifle ridiculous' ((104) p. 33). He summarizes his own view by comparing the economic system to 'a ball lying on a grassy slope. It might take quite a hard kick to move it. But, once moved, it might go further, especially if the hill was steep, than an initial kick of equal force would have moved it if it had been lying in a flat field. It might go the whole way down a mountainside' ((104) p. 32).

It should be clearly understood that the Second Harrod Problem is logically independent of the First. As Hahn and Matthews have commented:

> 'it is important to distinguish clearly between the two quite separate obstacles to steady growth that were considered by Harrod in his pioneering contribution. (1) The warranted rate may be unequal to the natural rate. (2) The warranted rate may itself be unstable, even without reference to the natural rate. The second of these problems is the "knife-edge" properly so-called, though the term is sometimes used confusingly to refer to the first problem as well[20] (Hahn and Matthews (85) p. 27).

Thus, any analysis which weakens the force of one problem does not necessarily weaken the other and it will be necessary to re-emphasize this point when discussing the so-called 'neoclassical' growth model in the next chapter. On the other hand, the two problems would, of course, interact in a real economic system. Consider, for example, a situation in which $G_A = G_W = G_n$ and steady-state growth at full employment is proceeding at the natural rate. If, as a result of an increase in s, the warranted rate increases, it will now be above the natural rate. The actual rate, G_A, must now diverge from the warranted rate because, as we have already noted, it cannot exceed the natural rate for long. If the actual rate is less than the warranted then v must exceed v_r and entrepreneurs will be induced to decrease investment, further reducing the actual rate and the economy is consequently dragged into a recession. In this situation, the First and Second Harrod problems are *combining* to generate a recession and unemployment. But the two *problems* are logically separate.

The Harrod Model: A Summary

The Harrod model of economic growth highlights the necessary conditions for full-employment steady growth in a developed capitalist economy. Three central issues have been noted:

(a) The *Possibility* of steady-state growth at full employment.
(b) The *Improbability* of steady-state growth at full employment.
(c) The *Instability* of the warranted rate of growth.

3.5 The Constant Capital–Output Ratio

Until now we have tended to interpret the fixed-proportions characteristic of the Harrod model as a reflection of the fixity of the technology and this interpretation is consistent with much of the retrospective literature on the Harrod model. If, however, the Harrod model *is* based upon the assumption that capital and labour simply cannot, in any circumstances, be substituted for each other in producing a given flow of output, then some of its central conclusions, in particular the First Harrod Problem, are extremely vulnerable to changes in specification. Solow, in his 1956 paper (244), which forms the basis of our discussion of simple neoclassical models of growth in the next chapter, argued that:

> 'this fundamental opposition of warranted and natural rates turns out in the end to flow from the crucial assumption that production takes place

[20] Some writers, notably Kregel, (see, for example, Kregel (152) p. 43) use the term 'knife edge' in a slightly different sense.

under conditions of fixed proportions. There is no possibility of substituting labour for capital in production. If this assumption is abandoned, the knife-edge notion of unstable balance seems to go with it'[21] (Solow (244) p. 65).

It is, however, the purpose of this section to suggest[22] that to interpret the fixed capital–output ratio in the Harrod model as though it were a simple property of the technology is to blur one of the central features of his analysis.

We can begin to see the nature of Harrod's argument by recalling the simple marginal productivity theory discussed in Chapter 2 (2.4(iv)) and noting that Harrod consistently maintained that he was asking: 'what behaviour of capital is required to be consistent with growth in other elements, *on the hypothesis that the rate of interest does not change*?' (emphasis in original—(100) pp. 21–2). Moreover, when discussing the assumption 'that the capital/income ratio is constant' he took pains to emphasize that this followed from the assumption that the rate of interest was constant.

Now, in the absence of risk and in competitive conditions, the rate of interest is equivalent to the rate of profit, and simple marginal productivity theory would suggest that the rate of profit is driven by competition to equal the marginal productivity of capital—i.e. the rate of change of output with respect to an incremental change in capital:

i.e. $$r = \frac{\Delta Y}{\Delta K}\left[\text{strictly } \frac{\partial Y}{\partial K} = r\right]$$

where ΔY and ΔK represent the increments in output and capital respectively. Referring to our definition of v in terms of the incremental capital–output ratio, $\Delta K/\Delta Y$, (equation (3.2.3)) it should be clear that a constant capital–output ratio implies a constant rate of interest and vice versa[23]. Harrod makes this inverse relationship between the rate of interest and the capital–output ratio more explicit in his 1960 paper (102) and comments that this 'should satisfy objectors who complain that I took no account of the substitutability of capital for other factors' ((102) p. 285).

Thus, we can see that *if* the rate of interest were fixed then v would be fixed and the constancy of the capital–output ratio would stem from an economic mechanism rather than from an arbitrary technological assump-

[21] It is interesting to note that Solow appears to be exhibiting the confusion between the two Harrod problems against which Hahn and Matthews later warned. As we shall see in the next chapter, Solow's 1956 approach removes one of the Harrod problems but the other remains (or is by-passed).

[22] Following Eisner (63). See also Burmeister and Dobell (34) p. 41.

[23] It is important to recall that we are assuming the absence of technical progress.

tion. Variations in v would require variations in the rate of interest and, in particular, Harrod concluded that

> 'fundamental conditions might require a steadily falling rate of interest. We found great difficulties in envisaging how the capital market could ever succeed in providing such a steady decline' (Harrod (100), p. 75).

Any force preventing the rate of interest from moving freely—such as Keynes' famous 'liquidity trap'[24]—would prevent the capital–output from moving freely and the 'Harrod Problems' could emerge even though v was technologically variable.

We can distinguish at least four approaches to the fixity of the capital–output ratio which would be consistent with Harrod's principal conclusions:

(a) The capital–output ratio is fixed as a *consequence of the technology*. Harrod's model has often been interpreted in this way although there seems little reason to believe that this is what he meant.

(b) The capital–output ratio is able to vary somewhat but not enough to *guarantee* the necessary equality of warranted and natural rates of growth.

(c) Capital and labour are technically substitutable but in practice v it fixed due to the *inflexibility of factor prices*, in particular the *interest rate*. This interpretation seems closest to what Harrod actually wrote and it could be interpreted as being a fundamentally Keynesian approach.

(d) The long-run rate of interest is determined by welfare requirements (see Chapter 9) and there is no reason for it to adjust to the level corresponding to the appropriate value for v in the equality of the warranted and natural rates. This is one interpretation of the position taken by Harrod in his 1960 essay (102) and, once again, emphasis is placed upon the possible degree of variability of the rate of interest.

To sum up: it would seem inappropriate to interpret Harrod as referring to a world in which capital and labour simply cannot be substituted for one another and the capital–output ratio is arbitrarily assumed to be a technological constant. It seems clear from Harrod's writings that he was fully aware of technological substitution possibilities but considered that an analysis based upon the assumption of a constant capital–output ratio was fruitful if the money and capital markets were not capable of producing variations in the rate of interest sufficient to ensure the existence and attainment of a 'Golden Age'. The Harrod problems can emerge even if v

[24] Some recent studies have contested whether Keynes actually included the idea of a liquidity trap, particularly in its textbook form, in *The General Theory*. See Leijonhufvud (158), or, for a briefer and more colourful exposition of the same point, Leijonhufvud (159).

is technically capable of variation but is constrained by a relatively inflexible interest rate. We shall have occasion to return to this issue when we discuss the so-called 'neoclassical' models of growth in Section 4.5.

3.6 Domar's Model of Growth

Domar's growth model ((57) and (58)) is frequently bracketed with that of Harrod because of the similarity between his central result and Harrod's 'fundamental equation'. This is rather to be regretted as the two theories are most fruitfully studied in their own right rather than as some kind of amalgam.

Domar's approach is to focus clearly on the *dual* nature of the rate of investment in a capitalist economy:

(a) Investment determines the actual level of income through the Keynesian multiplier process.

(b) Investment, by increasing the size of the capital stock (we are continuing to assume the absence of depreciation) increases the *maximum potential* level of income.

Domar's model can be easily formulated by exactly analagous steps to that of Harrod but some insight is gained if the game is played along the lines that Domar originally intended[25]. Let:

Y = the actual level of income or output.
$\bar{Y}$ = the *maximum potential* level of national income or output.
s = the constant average and marginal propensity to save.
I = the flow of investment.
σ = the 'potential social average investment productivity'.

Only σ of the above symbols requires any comment and it is fair to point out that this is a somewhat confusing notion to grasp. Domar defines it as:

$$\sigma = \frac{\dot{\bar{Y}}}{I}$$

Thus, σ refers to the rate of change of the *potential* capacity for the production of output associated with a given level of investment and we note for future reference that 'it does not imply that factors of production other than capital and technology remain constant' ((59) p. 74). Domar assumes that σ is a constant and, as a consequence, equation (3.6.1):

$$\dot{\bar{Y}} = \sigma I \qquad (3.6.1)$$

[25] The following constitutes an extremely condensed version of Domar's model and, in particular, the 'junking' process is ignored. The interested student is urged to consult the references: (57), (58).

is a comprehensive description of the supply side of the economy.

The actual level of income at any point in time is determined by a conventional simple multiplier process:

$$Y = \frac{1}{s} I$$

or, in terms of the rate of change of income,

$$\dot{Y} = \frac{1}{s} \dot{I} \qquad (3.6.2)$$

Let us assume, with Domar, that the economy is initially in a position of full-employment equilibrium which implies that $Y = \bar{Y}$. Domar's basic purpose was the discovery of the growth rate of investment that would *maintain* Y equal to $\bar{Y}$. Clearly, if Y *is* to remain equal to $\bar{Y}$ then $\dot{Y}$ must equal $\dot{\bar{Y}}$ and, combining equations (3.6.1) and (3.6.2), we obtain

$$\sigma I = \frac{1}{s} \dot{I}$$

or

$$\frac{\dot{I}}{I} = \sigma s \qquad (3.6.3)$$

Equation (3.6.3) demonstrates that, as both s and σ are assumed to be constant, the rate of growth of investment that will maintain actual income equal to the maximum potential level of income is a *constant* proportional rate of $s\sigma$. The astute student will already recognize the similarity between equation (3.6.3) and Harrod's 'fundamental equation'. Moreover, a little algebraic manipulation will confirm that the other major macroeconomic variables must grow at the same rate if equilibrium is to be maintained. For our purposes, the above exposition contains the core of Domar's theory and the similarities and differences between it and Harrod's model are discussed in the next section.

3.7 Harrod and Domar: Some Comparisons

The similarities between the central results of the models of economic growth associated with Harrod and Domar resulted in the joint title 'Harrod–Domar' being used to refer to their approach to growth. For Harrod, steady-state equilibrium growth necessitates that:

$$G_A = \frac{s}{v_r}$$

and, in particular, that $\dot{I}/I = s/v_r$[26]. Domar's model suggests that dynamic equilibrium requires that $\dot{I}/I$, the rate of growth of investment, must equal $s\sigma$. Now, σ is the potential increase in output per unit of investment while v_r is the number of units of new investment required to produce an extra unit of output. Both concepts assume that the new investment is properly utilized. (See Harrod (101) p. 452.) Thus, it is clear that $\sigma = 1/v$ and substitution demonstrates that the two equations are *formally identical.*

A 'perplexity' (Harrod (101)) does, however, arise. For Harrod, steady-state growth *at full employment* requires that the warranted rate, s/v_r, equals the natural rate, n. Domar, however, makes it clear that $s\sigma$ is, to him, the equilibrium rate of growth that will ensure 'the maintenance of full employment' *without any reference* to the rate of growth of the labour force. Now, Domar makes no explicit assumptions concerning the rate of growth of the labour force and it is therefore difficult to be certain of the exact relationship between Harrod's warranted and natural rates and Domar's equilibrium growth rate. Harrod ((101) pp. 452–6) sees the resolution of the difficulty in the all-encompassing nature of Domar's definition of the potential social average investment productivity which seems to imply that $s\sigma$ incorporates elements of both the warranted and natural rates. Domar asserts that σ will normally be below $\bar{\sigma}$, the absolute maximum potential average investment productivity. This is because (amongst other things) of 'the growth of other factors, such as labour' (Domar (59) p. 74). Thus, Harrod's interpretation is that $s\sigma$ really corresponds to his *natural* rate simply because σ is constrained from equalling $\bar{\sigma}$ by the rate of growth of the labour force. Harrod's view is therefore that

> 'If only there was enough labour, etc., the economy could grow as $s\bar{\sigma}$, which is really my "warranted" rate; but the lack of labour, etc., compresses potential growth to $s\sigma$, my "natural" rate' (Harrod (101) p. 456—notation slightly changed).

There is, however, a more fundamental difference between the approaches of Harrod and Domar. Notice that the Domar model does not include an investment function. His equilibrium growth rate is the rate that would ensure the continuing equality of actual and potential income *but the actual level of investment is not determined within the confines of the model.* Domar's approach is fundamentally one of deriving dynamic consistency conditions rather than of generating a *theory* of growth. Harrod, on the other hand, includes a specific investment function in the form of the simple accelerator mechanism.

[26] Harrod's own expedition never explicitly refers to the rate of growth of investment and, in his comments on Domar's model (101), he points out that this was a deliberate choice on his part in an effort to attain the greatest possible generality. It is, however, a consequence of our particular version of Harrod's model that the rate of growth of investment will equal s/v_r.

We can summarize the similarities and differences between the models of Harrod and Domar as follows:

(a) Both models begin from a fundamentally Keynesian framework but move into the long run by eschewing Keynes' assumption that the rate of investment did not increase the size of the capital stock. (An assumption which is clearly only suitable in the framework of the short-run analysis of income determination.)

(b) Both models generate an equilibrium condition implying a *constant* proportional rate of growth of the economy.

(c) Both models imply long-run difficulties in attaining equilibrium growth at full employment. In Harrod's model this problem arises because there is no mechanism to ensure the necessary equality of the warranted and natural rates and, moreover, the warranted rate of growth is inherently unstable. In the Domar model a similar problem arises because of what he sees as a chronic tendency to underinvest[27]. so that the rate of growth of investment does not, in general, equal $s\sigma$.

(d) Both models assume the equivalent of a constant capital–output ratio. Domar sees this as a convenient assumption as to the fixity of the technology. Harrod, on the other hand, argues from a fundamentally Keynesian scepticism as to the magnitude of possible variations in the interest rate.

(e) Both models involve an element of instability although the actual mechanism is much clearer and, perhaps, more fundamental in Harrod's model. Instability in the Harrod model stems, as we have seen, from the interaction of the investment function and the fundamental equation with entrepreneurial expectations. In the Domar model, investment incentives are continually weakened although the exact mechanism does not seem to be very clear.

(f) The 'vision' (see Leijonhufvud (158) pp. 9–11) of both writers is similar. Both visualize, as a plausible scenario, a long-run state of depression with chronic unemployment and idle capacity.

3.8 Harrod and Domar: Assessment and Conclusion

The proliferation of articles in economics journals on theories of economic growth has been one of the most notable features of the economic analysis of the past quarter of a century or so—and it is clear that Harrod and Domar were the pioneers of this movement. An allusion to, or sometimes an exposition of, their theories seems to have become mandatory in all but the most elementary macroeconomics textbooks, and to many undergraduates Harrod's theory *is* the modern theory of economic growth.

[27] This, Domar believes, stems from his 'junking' mechanism which we do not discuss.

$$K_1/K = \frac{v_1 Y_1}{v_1 Y_1 + v_2 Y_2} = \frac{v_1 I}{v_1 Y_1 + v_2 Y_2}$$

If, as in the Harrod model, $v_1 = v_2 = v$ then we can write

$$\mu = \frac{vI}{v(Y_1 + Y_2)} = \frac{vI}{vY} = \frac{I}{Y} = \frac{S}{Y} = s$$

Thus, we have shown that if the capital–output ratio in both categories is the same then $\mu = s$ and Fel'dman's rate of growth, μ/v_1 is equal to Harrod's warranted rate, s/v[26]. Fel'dman's model can easily be extended in a number of directions. The inclusion of depreciation poses no formal problems (see Domar (59) or Bose (28)) but it does suggest an interesting special case which is absent in our version of the model. It is possible that the initial capacity of the investment goods industry is sufficient only to replace worn-out equipment in both sectors. In such a situation growth can only take place as a result of an accidental or deliberate failure to replace worn-out assets in the consumption goods industry. In this situation, which bears a family resemblance to the 'low-level equilibrium trap' discussed in Chapter 4 (see 4.6(f) Case 2), a contraction of consumption output is necessary to begin a 'take-off into self-sustained growth' (see Rostow (216)). If a contraction of consumption output is impossible for subsistence reasons we can see that a special version of the 'vicious circle of poverty' theme will emerge.

5.8 Conclusion

We have discussed some of the general problems associated with two-sector models of a growing economy and have examined two particular models as examples of this class of theorizing. The construction of two-sector models is, however, almost an industry in itself and the interested student would only need to consult the structured reading list to gain some impression of the great varitey of assumptions (and often of conclusions) that can be employed in the manipulation of this type of model. On the other hand, it might be suggested that the analysis of two-sector models is now experiencing diminishing returns in that the more sophisticated models add only marginally to the insights gained from the simple models and many of the most important lessons were, of course, embodied in Volume 2 of Marx's *Capital* published almost a century ago.

[26] In Harrod's model the propensity to save, s, is determined by the individual preferences of the households of the economy whereas μ in the Fel'dman model is determined by the choice of the socialist government. Thus, despite the similarity of s/v and μ/v_1, the two growth rates bear quite different interpretations.

chapter 6
The Cambridge Critics

6.1 Introduction

For the past twenty[1] years or so, the neoclassical 'vision' of economic growth, together with many of the specific concepts and methods employed in its elaboration, has been subjected to a series of more or less successful attacks from a group of distinguished economic theorists collectively known as the 'Cambridge School' because of their association with the Faculty of Economics at the University of Cambridge, England. The principal defence has been mounted by a group of no less distinguished economists at the Massachusetts Institute of Technology (M.I.T) in Cambridge, Massachusetts. Few controversies in the history of economic thought have been conducted with so much vigour and, at times, virulence as the series of interconnected debates between the two Cambridges on the concept of 'capital' and the process of economic growth and technical change. Outside the relatively small circle of participants the debate has generated a mixture of emotions and conclusions—many being excited, others bewildered and a large number resolving on indifference[2]. There can be little doubt, however, that *some* of the problems around which the controversies have centred *are* important. In Harcourt's words: '... important issues—growth, distribution, accumulation, in fact all the classical, if not classic, puzzles of our trade—are being discussed' ((94) p. 14).

The modern[3] Cambridge School of economists revolves around the works of Professor D. G. Champernowne, Professor Lord Kahn, Professor Nicholas Kaldor, Dr Luigi Pasinetti, Professor Joan Robinson and Piero Sraffa—although Champernowne, Kahn and Sraffa have not

segue dopo cap. 4

[1] It could be argued that the debate dates back to the original controversies over Keynes' *General Theory*. Moreover, the principal propositions of one of the central contributions to the modern debate, Sraffa's (256), were formulated in the late 1920s. See Sraffa (256) p. vi.

[2] Harcourt has commented that: 'outside the two Cambridges these discussions have been regarded as "a little silly". How can grown men (and women) get so cross over matters like these?' (Harcourt (94) p. 119).

[3] The University of Cambridge has a lengthy tradition in economic theorizing. Malthus, Marshall, Pigou, Lord Keynes and Sir Dennis Robertson were all members of the University and the description 'Cambridge School' has, at various times, been applied to individuals and ideas very different to those discussed in this chapter.

have not emerged in the post-war world because of the systematic application of Keynesian policies of economic stabilization[29]. Despite these criticisms, the various facets of the Harrod–Domar achievement seem clear. They reintroduced the *idea of growth* to economic theory and the central role of the concepts of steady state and balanced growth, although foreshadowed in Cassel's exposition (36) of the 'regularly progressing state', stems from their work. They re-emphasized the 'classical' role of saving as the accumulation of capital after the Keynesian controversies on the role of saving in the context of the determination of the level of national income. The central ideas of their models were explicitly introduced into the study of economic development[30] and played a prominent part in the intellectual substructure of the remarkable thesis, due to Rostow (216), which has caused so much controversy in the study of the economic history of the growth of economies over the past two centuries[31]. Harrod's model was capable of being interpreted as a simultaneous explanation of trade cycles as well as economic growth. The influence of the Harrod–Domar approach has been so widespread that it simply cannot be encapsulated in a few short sentences—it should be viewed as being pervasive in almost every chapter of this book. However, with the advantage of hindsight, it seems that Harrod's most important contribution was his emphasis on the role of entrepreneurial expectations as a central source of difficulty for the attainment of full-employment steady-state growth—and this is a theme that re-emerges in the next chapter.

[29] But see the controversial paper by Matthews (177).

[30] Thus, for example, Higgins, in his famous textbook on economic development, writes: 'Harrod himself did not apply his system to the problems of underdevelopment, but that fact alone should not prevent us from doing so ((109) p. 113). For a partial antidote to the kind of enthusiasm exemplified by Higgins see Hirschman (111) Ch. 2.

[31] Rostow's concept of the 'take-off into self-sustained growth' is premised on a 'dynamic theory of production' which seems to owe a great deal to the influence of the Harrod–Domar line of thought. See, for example, p. 37 of (216).

chapter 4
The Neoclassical One-sector Model of Growth

4.1 Introduction

Despite the seminal roles of Harrod and Domar in reawakening interest in the problems of growth and long-run accumulation, the so-called 'neoclassical' approach to the analysis of a growing economy has attracted substantially more professional interest, and even enthusiasm, during the past fifteen years or so—such that it can now be said to represent the *dominant* method of growth economics. As with all general categorizations of individuals and ideas, it is difficult to be precise as to exactly what is meant by the description 'neoclassical'—and this lack of precision serves as a fertile source of confusion for the student. An unambiguous definition of neoclassical economics is simply not possible—but we can try and identify three separate strands of thought which contribute towards the modern conception of the term.

(i) The original neoclassical economists were those[1] who, in the latter part of the nineteenth century, using the concepts of the 'marginalist revolution' (see Blaug (25) Ch. 8)—marginal utility and marginal productivity—concentrated their attention on the analysis of the pricing of individual goods and factors of production in competitive markets, and on the possible existence of a *set* of prices which would ensure the equality of supply and demand in *all* the markets of the economy[2]. Thus, one possible modern interpretation of the description 'neoclassical' is that it refers to that body of economic theory which incorporates some of the central ideas of the nineteenth century neoclassicals whether through a general, 'rational', maximizing, microeconomic approach to economic phenomena, or through the use of specific theories and concepts such as a marginal productivity explanation of wages (see 2.4(b)(iv)) or ideas of perfect competition and perfect flexibility of all prices[3].

[1] e.g. Marshall, Jevons, Walras, Wicksell, Pareto, Clark, Edgeworth and Fisher. A particularly stimulating introduction to this aspect of economic thought is available in Samuelson (227) Ch. 42. More details can be obtained from Blaug, (25), particularly Chs 8, 9, and 11, and the references contained therein.

[2] The powerful idea of 'general economic equilibrium' is usually attributed to Walras. See Arrow and Hahn (14) Ch. 1 for a succinct description of the origin and development of the idea.

[3] A brief paragraph, purporting to summarize a multitude of ideas originating from a host of writers, is inevitably somewhat misleading. The interested reader is urged to consult the references of footnote 1.

(ii) Karl Marx invented the description 'classical' economics to refer to the writings of Ricardo, Mill and their *predecessors*, but Keynes, in his *General Theory* (see (140) p. 3 footnote 1) although admitting that he was perhaps 'perpetrating a solecism', used the term to cover all the economists who had *followed* Ricardo and who had believed that there were forces in the economy that guaranteed the generation of a *full-employment* level of income[4]. With the gradual acceptance of the 'Keynesian Revolution' (see Klein (144)) many economists, notably Samuelson, argued that a 'neoclassical synthesis' was possible in the sense that, once the validity of Keynes' theory was recognized, and governments could take action via fiscal and monetary policies to maintain full-employment income, then the microeconomic theories of the 'classical' (in Keynes' terms) economists could once more be fruitfully used[5]. Thus, a *second* possible meaning of the term 'neoclassical' is the description of theories which, while not necessarily denying the validity of Keynes' strictures, ignore what are often called Keynesian 'difficulties' by assuming the existence of a government which persistently, continually and successfully manipulates the policy instruments at its disposal so as to maintain the full-employment level of aggregate demand. Meade's famous exposition of a neoclassical theory of growth is therefore 'based on the assumption of an ideally successful Keynesian policy which at every point of time manages to keep the value of investment at the desired level' although he asserts that his theory is 'certainly not classical in the sense of being pre-Keynesian'. (See Meade (178) p. ix.)

(iii) The original 'classical' economists of the first half of the nineteenth century were much more concerned with the *long-run* forces that governed the macroeconomy than with the behaviour of individual markets—which was the focus of attention of the nineteenth-century neoclassicals. (See (i) above.) *A third* interpretation of the term 'neoclassical' would therefore emphasize the subordination of short-term problems to long-term trends in modern neoclassical economic theory.

The three factors outlined above provide only a rough sketch of the multifarious meanings that can be, and have been, applied to the term 'neoclassical' in the context of economic theory. To many of the opponents of neoclassical economics it is a doctrine, or even a dogma, which, at least in part, is concerned with the justification of the central features of a capitalist economy. To most of the adherents of neoclassical economics it is an apparatus of the mind or a technique of thinking—'an economic theory derived from some kind of rational behaviour' (Stiglitz and Uzawa (258) p. 310) which can, in many circumstances, produce fruitful insights into economic processes and problems. Whether the description 'neoclassical' is interpreted as an accolade or an insult seems to depend at least as much on one's philosophical and political outlook as on the details of the technical assumptions one employs in developing an economic theory[6].

Why has the neoclassical approach to growth theory dominated contemporary discussion? We can outline a number of *possibilities*.

(i) *The conclusions of neoclassical growth theory accord better with the 'facts' of real growth experience than any alternative formulation.*

Friedman's methodological standpoint (see Ch. 1) would imply that this would be a primary test of the value of any theory. Unfortunately, as we will see, the predictions of simple neoclassical growth theory which are clearly different from those of other theories are not particularly amenable to empirical testing. Although the neoclassical model can generate conclusions which conform roughly with most of Kaldor's 'stylized facts' (see 1.3), other theories are capable of a similar achievement. It would be difficult to maintain that the dominance of neoclassical growth theory stems from any decisive evidence of its explanatory or predictive properties.

(ii) *The majority of professional economists have been brought up in the neoclassical tradition and a neoclassical framework is therefore more 'natural' to them—both in developing their own theories and in assessing the work of others.*

There can be little doubt of the difficulty of changing any ingrained habit of thought or thinking. Keynes spoke of his 'long struggle of escape' ((140) p. viii) from the conventional approach to the analysis of unemployment. Given that the neoclassical approach dominates the teaching of economics in the Western world, it would be surprising if a non-neoclassical growth theory attained any large measure of acceptance independently of a transformation in economic theory in general[7].

(4) Keynes has often been accused of labelling every economist prior to himself as 'classical'. Moreover, his characterization of all 'classical' economists as believing in the automaticity of full employment lays him open to the charge of erecting a 'straw man'. It seems indisputable, however, that, prior to Keynes' contribution, economists were not, in general, concerned with the determinants of the level of aggregate income.

(5) Interestingly, the phrase 'neoclassical synthesis' featured prominently in earlier editions of Samuelson's famous textbook (see, for example, p. 361 of the 6th edition) but has apparently been expunged from the latest edition. (See p. 372 of (227).)

(6) Harcourt ((94) p. 13) has suggested: 'that if one were told whether an economist was fundamentally sympathetic or hostile to basic capitalist institutions, especially private property and the related rights to income streams, or whether he were a hawk or a dove in his views on the Vietnam War, one could predict with a considerable degree of accuracy his general approach to economic theory'. There is probably an element of truth in this statement—though it appears to confuse *correlation* with *causality*.

(7) It might be argued that Keynes' insights have become part of the 'natural' framework of thought of the majority of professional economists—as exemplified in the phrase 'We're all Keynesians now'. On the other hand, it is not clear that Keynes' 'vision' of the economy, as opposed to a neoclassical formalization of some of his major views, has ever been fully accepted. See the writers on the 'new' view of Keynes: (45), (158) and (110).

(iii) *Neoclassical growth theory is particularly tractable, while remaining, for many, intellectually satisfying, and it is suitable for extension in a number of different directions.*

One of the central attractions of the neoclassical approach to growth is that it can be set out simply and clearly as a system of equations and easily manipulated to produce what appear to be unambiguous conclusions. The apparatus is easily adapted to accommodate different assumptions and even the most inexperienced practitioner quickly feels capable of producing his 'own' theorems on the process of growth[8].

Neoclassical models of growth can be said to stem from two papers, by Solow (244) and Swan (262), that were published in 1956 although most of the characteristics of the neoclassical approach were included in a paper by Tobin (266) published in the previous year. Our exposition will concentrate on Solow's famous paper as a useful example of the class of simple neoclassical models. We will not, however, duplicate exactly his argument or notation.

4.2 The Assumptions

Since Solow's avowed purpose was an examination, and demonstration of the special nature of the 'fundamental opposition of warranted and natural rates' (what we have called the First Harrod Problem—see 3.3) most of his paper is devoted to 'a model of long-run growth which accepts all the Harrod–Domar assumptions except that of fixed proportions' (Solow (244) pp. 65 and 66). We can therefore list these assumptions without much comment.

ASSUMPTION 4.2.1

Solow's theory is conducted in the context of a model economy in which only *one good* is produced: 'There is only one commodity, output as a whole, whose rate of production is designated $Y(t)$. Thus we can speak unambiguously of the community's real income' ((244) p. 66). It is therefore clear that we are in the realm of 'parable' (see 2.2(d)) in which a single all-purpose commodity is produced and either consumed or invested. Notice that this formation implies that no Keynesian distinction is drawn between those who save and those who invest—saving simply *is* investment and no separate investment function need be included in the model. The community's stock of capital, $K(t)$, 'takes the form of an accumulation of the composite commodity'. In other words, in terms of a 'corn' parable, any 'corn' not eaten is saved and *automatically* becomes part of the stock of 'corn' capital. This assumption is clearly very powerful. It removes all the aggregation difficulties discussed in Chapter 2 and precludes any problem of a discrepancy between ex-ante saving and investment resulting from a dichotomy between savers and investors.

[8] Advanced neoclassical models are not, of course, as simple as is suggested here but, as Samuelson has pointed out, there is a kind of Gresham's Law of economics whereby simple theories gain more attention (and thereby become identified as *the* version of the theory) than the difficult.

ASSUMPTION 4.2.2

As in the Harrod model, a *simple proportional savings function* is assumed:

$$S = sY \tag{4.2.1}$$

where $0 < s < 1$.

Solow employed this assumption so that the similarities and differences between his model and Harrod's would be highlighted. A more thorough-going neoclassical approach would attempt to derive the savings behaviour of the community from the intertemporal preferences of the individuals within the community and the incomes which they expect to receive during their lifetimes.

ASSUMPTION 4.2.3

The capital stock does not depreciate. This assumption is not necessary for the development of the model and is employed here for simplicity. Depreciation is introduced in 4.7 below. Investment is simply the rate of increase of the capital stock of the composite commodity:

$$\dot{K} = I \tag{4.2.2}$$

and, given that investment is *identically* equal to saving (from Assumption 4.2.1) then Assumption 4.2.2 implies that we can write

$$\dot{K} = S \tag{4.2.3}$$

or,

$$\dot{K} = sY \tag{4.2.4}$$

ASSUMPTION 4.2.4

The labour force grows at an exogenous constant proportional rate n. i.e.

$$\frac{\dot{L}}{L} = n \tag{4.2.5}$$

ASSUMPTION 4.2.5

The technical possibilities of the economy are represented by a *continuous, constant returns to scale, aggregate production function:*

$$Y = F(K, L) \tag{4.2.6}$$

The assumption of constant returns implies that equation (4.2.6) can be written in the *intensive* form discussed in Section 2.4 (p. 28).

$$y = f(k) \tag{4.2.7}$$

where $\quad y = Y/L$ and $k = K/L$.

Thus, equation (4.2.7) states that output per labourer is a function of capital per labourer. Solow argued that 'constant returns to scale seems the natural assumption to make in a theory of growth'. It is clear that the assumption of a continuous aggregate production function is fundamentally different to the fixed-coefficients form used by Harrod and Domar.

We will, moreover, assume that the aggregate production function satisfies the following conditions:

(i) The marginal product of capital, denoted $f'(k)$ is positive for all levels of the capital–labour ratio. (See Assumption 2.4.1.)

i.e. $\quad f'(k) > 0$ for all k.

(ii) The marginal product of capital diminishes as capital per labourer increases. (See Assumption 2.4.2.)

● Mathematically, this implies that $f''(k) < 0$ for all k. ●

(iii) As the ratio of capital to labour, k, tends towards infinity (i.e. as it becomes larger and larger), the marginal product of capital tends towards zero. At very high levels of the capital–labour ratio, the marginal product of capital becomes very small.

● In mathematical terms, we can write this condition as:

$$\lim_{k \to \infty} f'(k) = 0$$ ●

(iv) As the ratio of capital to labour, k, tends towards zero, the marginal product of capital tends towards infinity.

● This condition may be written mathematically as:

$$\lim_{k \to 0} f'(k) = \infty$$ ●

(v) No output can be produced without any capital.

i.e. $\quad f(0) = 0$

(vi) An indefinitely high level of output per labourer is associated with an indefinitely large ratio of capital to labour.

i.e. $\quad f(\infty) = \infty$

An aggregate production function satisfying the above conditions (often known as the Inada conditions—see Inada (113)) is usually referred to as WELL-BEHAVED. Three of the conditions ((i), (ii) and (v)) are included in Assumptions 2.4.1–4 and a diagram of a per-worker production function which satisfies all six conditions is available as Fig. 2.4. Solow did not assume that the aggregate production function was well-behaved but, having demonstrated the power of the assumption of well-behavedness, it will be possible to examine the more general case of an unrestricted production function later in the chapter.

Assumptions 4.2.1–5 are the *explicit* assumptions of the Solow model. It is, however, important to realize that Solow's model presents 'the neoclassical side of the coin' ((244) p. 91) and that less obvious neoclassical assumptions permeate his treatment. The mechanism whereby most Keynesian 'difficulties' have been 'shunted aside' is the necessary identity of ex-ante saving and ex-ante investment (as a consequence of Assumption 4.2.1) with the resulting 'absence of an investment function and the consequent failure to assign a major role to entrepreneurial *expectations* about the future' (Sen (237) p. 23—emphasis added). Moreover, the factor markets in the Solow model work perfectly—with the wage rate and rate of profit adjusting smoothly and instantaneously to changing circumstances.

4.3 The Fundamental Equation of Neoclassical Economic Growth

Harrod's model of economic growth revolved around a 'fundamental equation' defined in terms of the rate of growth of output. The neoclassical model implies a *different* fundamental equation which traces the path of the *capital–labour* ratio through time. As we will see, (Section 4.4) a knowledge of whether the capital–labour ratio, k, is growing, declining or staying constant is sufficient, in the context of the neoclassical model, to characterize the growth rates of all the major macroeconomic variables. Given assumptions 4.2.1–5 it is a relatively simple matter to derive this fundamental equation of a neoclassical one-good model of growth. We proceed in four stages.

(a) In the one-good economy, income, measured unambiguously in terms of the single all-purpose good, is identically equal to aggregate consumption plus aggregate investment[9]:

$$Y \equiv C + I \tag{4.3.1}$$

where $Y =$ income, $C =$ consumption and $I =$ investment.

[9] In the absence of government expenditure and foreign trade. Notice that, in a conventional Keynesian macroeconomic model, it would be essential to distinguish between writing $Y = C + I$ as an ex-ante equilibrium condition and $Y \equiv C + I$ as an ex-post accounting identity. (See, for example, Rowan (218) Ch. 8.) This is not necessary in the one-good neoclassical model simply because whatever is not consumed is automatically invested to become part of the 'corn' capital stock.

We can transform equation (4.3.1) into per-worker terms by dividing through by L:

$$\frac{Y}{L} \equiv \frac{C}{L} + \frac{I}{L} \tag{4.3.2}$$

More precisely,

$$\frac{Y}{L}(t) \equiv \frac{C}{L}(t) + \frac{I}{L}(t) \tag{4.3.3}$$

i.e. output per worker at time t is divided between consumption per worker at time t and investment per worker at time t.

From Assumption 4.2.5. we know that output per worker, $Y/L = y$, is a function of capital per worker, so that equation (4.3.3) can be written as

$$f[k(t)] = \frac{C}{L}(t) + \frac{I}{L}(t) \tag{4.3.4}$$

(b) Consider the capital–labour ratio $k = K/L$. It is clear that if the capital stock, K, and the labour force, L, are both growing at the *same* rate then the rate of growth of k will be zero—i.e. it will remain unchanged. If the proportionate rate of growth of K, which, for convenience, we will denote as $\hat{K}$ $(= \dot{K}/K)$, is greater than the proportionate rate of growth of the labour force, $\hat{L}$ $(= \dot{L}/L)$, then clearly the capital–labour ratio will be growing, i.e. $\hat{k} = \dot{k}/k > 0$. Similarly, if the rate of growth of K, $\hat{K}$, is *less than* the rate of growth of L, $\hat{L}$, then k will be declining—i.e. $\hat{k} = \dot{k}/k < 0$. A little reflection[10] on these comments will confirm that the rate of growth of the capital–labour ratio must equal the rate of growth of the capital stock minus the rate of growth of the labour force:

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L} \tag{4.3.5}$$

or,

$$\hat{k} = \hat{K} - \hat{L}$$

By Assumption 4.2.4 we know that $\hat{L}$, the rate of growth of the labour force, is a constant n—so that equation (4.3.5) can be written

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - n$$

[10] The mathematician will recognize that by taking natural logarithms of $k = K/L$ we obtain: $\ln k = \ln K - \ln L$, where ln denotes the natural logarithm. By logarithmic differentiation (see Allen (6) Ch. X) we obtain:

$$\frac{1}{k} \cdot \frac{dk}{dt} = \frac{1}{K} \cdot \frac{dK}{dt} - \frac{1}{L} \cdot \frac{dL}{dt}$$

Multiplying both sides by $k = K/L$ we obtain:

$$\dot{k} = \frac{\dot{K}}{K} \cdot \frac{K}{L} - n\frac{K}{L}$$

and, by cancelling the K's

$$\dot{k} = \frac{\dot{K}}{L} - nk \tag{4.3.6}$$

or, alternatively,

$$\frac{\dot{K}}{L} = \dot{k} + nk \tag{4.3.7}$$

(c) From Assumption 4.2.3 we know that $\dot{K} = I$. Thus $I/L = \dot{K}/L$. Equation (4.3.7) demonstrates that $\dot{K}/L = \dot{k} + nk$. We can therefore replace I/L by $\dot{k} + nk$ in equation (4.3.4) to obtain[11]:

$$\boxed{f(k) = \frac{C}{L} + \dot{k} + nk} \tag{4.3.8}$$

Equation (4.3.8) simply states that output per worker, $f(k)$, is allocated to three uses—consumption per worker, C/L, a portion of investment which maintains the capital–labour ratio constant in the face of a growing labour force, nk, and a portion of investment, $\dot{k}$, which actually increases[12] the capital–labour ratio. The process whereby the capital–labour ratio actually increases is often called 'capital deepening' whereas the process whereby capital goods are accumulated merely to match the growing labour force and keep the capital–labour ratio constant is referred to as 'capital widening'. (See Samuelson (227) p. 602*n*. and pp. 742–3.) Thus, in these terms equation (4.3.8) shows that output per worker is divided between consumption per worker, 'capital widening', and 'capital deepening'.

(d) Equation (4.3.8) can easily be transformed into what we will call the fundamental equation of neoclassical economic growth. Rearranging, we obtain

$$\dot{k} = f(k) - \frac{C}{L} - nk$$

[11] For clarity, the t's are omitted from equation (4.3.8).

[12] $\dot{k}$ can, of course, be negative in which case this term indicates that there are actual net reductions in the capital–labour ratio—capital is becoming 'shallower'.

which, recalling that $f(k) = y = Y/L$, can be written as

$$\dot{k} = \frac{Y}{L} - \frac{C}{L} - nk \qquad (4.3.9)$$

Now, in the context of the one-sector neoclassical model, the difference between output per worker, Y/L, and consumption per worker, C/L, is clearly savings per worker, S/L. Thus (4.3.9) can be written as

$$\dot{k} = \frac{S}{L} - nk \qquad (4.3.10)$$

Assumption 4.2.2 asserts that total saving, S, is proportional to total income: $S = sY$ so that equation (4.3.10) can be written as

$$\dot{k} = \frac{sY}{L} - nk$$

and, given that $Y/L = y = f(k)$ (Assumption 4.2.5), we obtain the fundamental equation that we require:

$$\boxed{\dot{k} = sf(k) - nk \qquad (4.3.11)}$$

This is the *fundamental equation of neoclassical economic growth* corresponding to equation (6) of Solow's original paper[13].

Consider the right-hand side of equation (4.3.11). We know that $sf(k)$ is simply saving per worker and, since in this model saving automatically becomes investment, it can also be interpreted as the flow of investment per worker. The second term, nk, is the amount of investment that would be required to keep the capital–labour ratio constant given that the labour force is growing at a constant proportionate rate n—i.e. 'capital widening'. Thus, devoid of all the manipulation required for its derivation, equation (4.3.11) is saying something rather obvious. The rate of change of the capital–labour ratio, $\dot{k}$, is determined by the difference between the amount of saving (and investment) per worker and the amount required to keep the capital–labour ratio constant as the labour force grows. If the saving per worker of the community is greater than this amount then it is clear that the capital stock will grow faster than the labour force and the capital–labour ratio will consequently increase.

4.4 Two Fundamental Propositions of Neoclassical Economic Growth

We can now use the fundamental equation to demonstrate two central propositions associated with the neoclassical model of economic growth.

[13] Solow uses the symbol r for the capital–labour ratio.

PROPOSITION 1

Given Assumptions 4.2.1–5, a balanced-growth (steady-state) solution for the model exists; this balanced growth solution is stable in the sense that, whatever the initial values of all the variables in the model, the economy moves steadily towards the balanced growth path.

This proposition is easily demonstrated with the aid of Fig. 4.1 which is not as complicated as it looks. Part A graphs the 'well-behaved' *intensive* production function, $f(k)$, as in Fig. 2.4. Each point along the curve $f(k)$ illustrates the quantity of output per worker, y, associated with any given

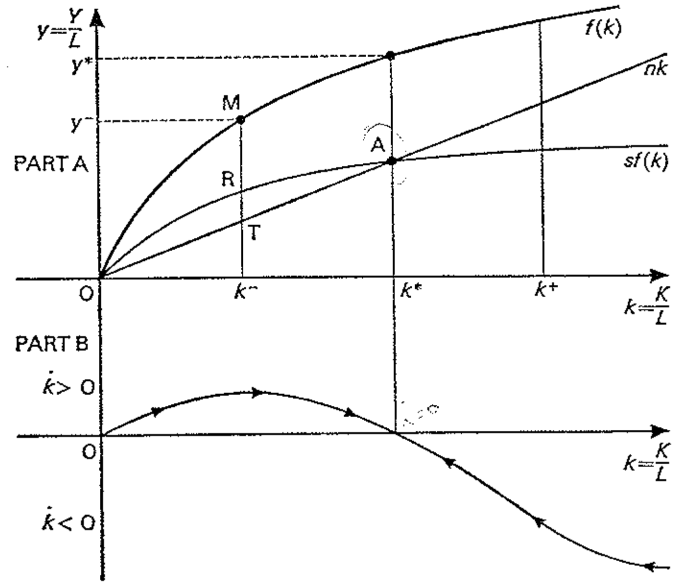

Fig. 4.1

level of capital per worker. Thus, for example, the capital–labour ratio of k^- in the diagram implies a flow of output per worker of y^-. A fraction s (the propensity to save) of any level of output per worker is saved and the curve $sf(k)$ therefore graphs the level of savings per worker associated with any level of the capital–labour ratio[14]. Thus, the distance k^-R measures saving (and investment) per worker and the distance MR measures

[14] The *shape* of the curve $sf(k)$ reflects the changes in the quantity of output per worker, $f(k)$, produced as the capital–labour ratio changes. The fraction of output per worker that is saved is, by assumption, a constant.

consumption per worker given a capital–labour ratio of k^-[15]. The line nk is drawn with its slope reflecting the *exogenous* constant proportional rate of growth of the labour force.

(a) We first demonstrate that the capital–labour ratio k^* (and the associated output per worker y^*) in the diagram implies a balanced growth path. At k^*, nk and $sf(k)$ intersect (Point A in Fig. 4.1) and are therefore equal. But, from the fundamental equation

$$\dot{k} = sf(k) - nk$$

it is clear that this equality implies that $\dot{k} = 0$, i.e. when savings per worker exactly match the quantity required to keep the growing labour force equipped then the rate of change, $\dot{k}$, of the capital–labour ratio will equal zero and the capital–labour ratio will remain at the constant level k^*.

Now, if $k = K/L$ is a *constant* and the labour force, L, is growing at a rate n then the capital stock must be growing at the same rate, i.e. constant k^* implies that $\dot{K}/K = n$. Similarly, the constant level of the capital–labour ratio, k^*, implies a constant level of output per worker, y^*. But if $y = Y/L$ is to remain constant with the labour force growing at the constant exogenous rate n then Y must be growing at the same rate, i.e. constant y^* implies that $\dot{Y}/Y = n$.

We have therefore demonstrated that, at the constant capital–labour ratio of k^*, all the relevant variables grow at the *same* constant rate—the rate of growth of the labour force or, in Harrod's terminology, the natural rate of growth. Such a situation accords with our definitions of steady state and balanced growth. For such a solution to *exist* then the line nk must intersect the curve $sf(k)$. For the solution to be non-trivial this intersection must occur at positive levels of output and capital per worker (i.e. the curves intersect at the origin but the implied zero level of output per worker is hardly economically meaningful). We will investigate the role of the assumption of 'well-behavedness' of the intensive production function in ensuring the existence of a solution in Sections 4.5 and 4.6.

(b) We have demonstrated that if the capital–labour ratio in our model economy *happened* to be k^* then balanced growth would occur—with output and capital stock growing at the constant proportional rate n, the rate of growth of the labour force. What, however, happens if the capital–labour ratio is greater, say k^+, or less, say k^-, than the critical level, k^*, of the capital–labour ratio which is associated with balanced growth of the economy?

Consider the capital–labour ratio k^-. Inspection of Fig. 4.1 reveals that, at this level of the capital–labour ratio, saving per worker, $sf(k)$ (measured by k^-R) is *greater* than nk (measured by k^-T). Consideration of the fundamental equation

$$\dot{k} = sf(k) - nk > 0$$

reveals that $\dot{k}$, the rate of change of the capital–labour ratio, must be positive and the capital–labour ratio must be *growing*. We can see what is happening most clearly by examining the inequality

$$sf(k) > nk \tag{4.4.1}$$

implied by the capital–labour ratio k^-. Now, an inequality will still hold if both sides are divided by the same positive quantity[16]. Thus, dividing both sides of (4.4.1) by k we obtain

$$\frac{sf(k)}{k} > n \tag{4.4.2}$$

Recalling that $f(k) = Y/L$ and $k = K/L$ (4.4.2), can be written as

$$s\frac{Y}{L}\cdot\frac{L}{K} > n$$

or, cancelling the L's

$$\frac{sY}{K} > n \tag{4.4.3}$$

But, in this neoclassical model, aggregate savings, $S = sY$, are automatically translated into investment, I, which, given the assumption of no depreciation, equals the rate of change of the capital stock, $\dot{K}$. Thus, (4.4.3) can be written as

$$\frac{\dot{K}}{K} > n \tag{4.4.4}$$

Now, $\dot{K}/K$ is the rate of growth of the capital stock and we have shown that, at the capital–labour ratio k^-, it must be greater than n, the rate of growth of the labour force. It is therefore clear that the capital–labour ratio must be *growing*—which is what we deduced directly from the fundamental equation. Put simply, a capital–labour ratio of k^-, combined with our assumptions, implies that the quantity of saving and investment

[15] The unit of measurement on the vertical axis of Part A of Fig. 4.1 is 'quantity of composite commodity (e.g. 'corn') per worker'. Output per worker, $f(k)$, saving per worker, $sf(k)$, and consumption per worker, $f(k) - sf(k)$, are all measured in the same units and they can therefore be simultaneously graphed on the same diagram.

[16] For the rules of manipulating inequalities see Parry Lewis (193) Ch. 1.

per worker is greater than that required merely to keep the growing labour force equipped at that capital–labour ratio. There is therefore a surplus of saving and investment per worker which implies that the capital–labour ratio *must* increase. This argument will, of course, apply to *all* values of the capital–labour ratio to the left of k^*. We have therefore demonstrated that for any level of the capital–labour ratio less than that required for balanced growth, there is a mechanism which is increasing the capital–labour ratio up to the required level.[17] Similarly, consider the capital–labour ratio k^+ which is greater than the balanced growth capital–labour ratio k^*. Given k^+, it is clear from the diagram that $sf(k) < nk$, i.e. savings per worker are insufficient to equip the growing labour force at that capital–labour ratio. From the fundamental equation

$$\dot{k} = sf(k) - nk$$

it is clear that $\dot{k}$ must be negative and the capital–labour ratio is *falling*. If we reversed the argument of inequalities 4.4.1–4 it would be clear that, given the capital–labour ratio k^+, the rate of growth of the capital stock is *less* than the rate of growth of the labour force which, as we deduced from the fundamental equation, means that the capital–labour ratio must fall. This argument applies to *any* capital–labour ratio to the right of k^*.

We have therefore established that, *whatever* the initial capital–labour ratio happens to be, a process of smooth convergence to balanced growth can be expected in this neoclassical economy. Part B of Fig. 4.1 highlights this process by graphing $\dot{k}$ against k. At all points to the left of k^* (except the origin) $\dot{k}$ is positive. At all points to the right of k^*, $\dot{k}$ is negative. The arrows on the curve show the direction in which the capital–labour ratio is moving. Part B is an example of what is called a PHASE DIAGRAM in the advanced literature on dynamic economics. This process may take a considerable time (see 4.7 (ii)) but the *long-run* prospect for this kind of neoclassical economy is *balanced growth at Harrod's natural rate of growth—the rate of growth of the labour force*. Once the capital–labour ratio k^* is attained, output and capital grow at the same constant proportional rate n, and output per worker, capital per worker, consumption per worker, and saving per worker all remain constant. This result differs radically from Harrod's 'vision' of economic growth and this is discussed in Section 4.5 below.

PROPOSITION 2

The balanced rate of growth in the neoclassical model is the constant exogenous rate of growth of the labour force. In the long run the economy converges to the balanced growth path. The long-run rate of growth of a neoclassical economy is therefore n, and is entirely independent of the proportion of income saved.

[17] Savings per worker equal nk at the origin and the capital–labour ratio is therefore not changing. But any positive level of the capital–labour ratio (i.e. any perturbation away from the origin) will tend to move towards the level k^*.

This neoclassical proposition seems paradoxical in the extreme. It seems to contradict the commonplace of the policy maker that an increase in the rate of economic growth requires more saving and investment. There is, as we will see from Fig. 4.2, a simple explanation. Imagine that the economy has a capital–labour ratio of k^* and, since $sf(k) = nk$, is therefore enjoying balanced growth at the rate n. What would be the effect of a large increase in the propensity to save, s, shifting the curve $sf(k)$ upwards to $s^+f(k)$? The exogenous rate of growth of the labour force remains

Cítale

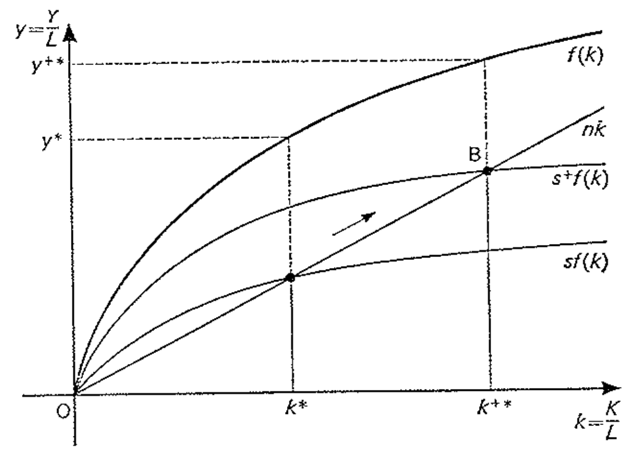

Fig. 4.2

unchanged so the new intersection of nk with the curve of savings per worker is at the point B in the diagram. Given the capital–labour ratio k^*, savings per worker with the new propensity to save are *greater* than required merely to keep the growing labour force equipped at that value of the capital–labour ratio. We can deduce directly from the fundamental equation (or with an identical argument to that implied by the inequalities (4.4.1)–(4.4.4)) that the capital–labour ratio must begin to rise, as indicated in the diagram by the arrow, and, as in the discussion of Proposition 1, it will continue to rise until k^{+*} is reached and a new balanced growth path is achieved. However, at k^{+*}, $s^+f(k)$ equals nk and the capital–labour ratio is again constant. With the labour force growing at the constant rate n, constancy of the capital–labour ratio implies that the capital stock and national output *must* be growing at the same constant rate n. (See Part (a) of Proposition 1.)

Thus, *despite* the increase in the propensity to save, the long-run rate of growth of income and the capital stock is still what it was prior to the increase in the savings rate. In the time that it takes to move from the original capital–labour ratio to the new balanced growth capital–labour ratio, the rate of growth of output and capital will temporarily increase[18] but the central conclusion of the neoclassical model of economic growth is that no *permanent* increase in the rates of growth of output and capital can be achieved by manipulating the propensity to save and invest of the economy. A policy of continually increasing the propensity to save so as to achieve a succession of 'temporary' increases in the rate of growth is limited by the fact that savings per worker cannot exceed output per worker and, assuming some minimum acceptable level of consumption per worker, the propensity to save must be less than one. The student will, however, notice that the increase in the propensity to save *does* increase the long-run *level* of output and income per worker to y^{+*}. *But the long-run rate of growth is totally unaffected by the propensity to save.*

We can therefore summarize the central conclusions of this simple neoclassical 'parable' of economic growth:

(1) The long-run rate of growth of the capital stock and national income is the rate of growth of the labour force which is assumed to be an exogenous constant n.

(2) The economy *invariably* tends to a balanced growth path whatever the initial capital–labour ratio.

(3) Output per worker, capital per worker, consumption per worker and saving per worker are all *constant* in the long run.

(4) Once and for all increases in the propensity to save, while increasing the levels of output per worker, y, and capital per worker, k, *do not* produce any change in the long-run rate of economic growth.

4.5 Harrod's Problems and the Neoclassical Model

In Chapter 3 it was demonstrated that Harrod's approach to the analysis of economic growth involved two fundamental problems:

(1) There was no reason why the warranted rate of growth, s/v, should equal the natural rate of growth n, and no mechanism whereby this equality could eventually be achieved.

(2) The warranted rate of growth, when combined with certain 'plausible' expectational patterns of entrepreneurs, could be intrinsically unstable.

It is clear that the principal conclusion of the neoclassical model of economic growth—smooth convergence to a path of balanced growth at the natural rate, n—completely *by-passes* these problems. Let us consider the means whereby this more optimistic 'vision' is achieved.

(a) We have already shown that balanced growth in the neoclassical model implies that

$$sf(k) = nk$$

or, dividing both sides of the equation by k,

$$\frac{sf(k)}{k} = n \tag{4.5.1}$$

Substituting Y/L for $f(k)$ and K/L for k, (4.5.1) can be written as

$$s\frac{Y}{L}\cdot\frac{L}{K} = n$$

or

$$s\frac{Y}{K} = n \tag{4.5.2}$$

But, from the definition of the capital–output ratio, $Y/K = 1/v$. Hence, (4.5.2) implies that

$$\frac{s}{v} = n$$

or, in Harrod's terms, the balanced growth capital–labour ratio, k^*, and output per worker, y^*, in Fig. 4.1, together imply that the warranted rate of growth equals the natural rate. We have already seen (Proposition 1) that the economy gravitates steadily to the balanced growth path in the neoclassical model; so it seems clear that the 'fundamental opposition of warranted and natural rates' is no longer a problem in the neoclassical model. The reasons are simple. Whereas in Harrod's model s, v, and n were all fixed constants the neoclassical assumption of a continuous aggregate production function implies that there exist a *whole spectrum* of values of the capital–output ratio and the economy adjusts to that particular value which ensures that the warranted rate equals the natural rate. Figure 4.3 illustrates the preceding comments.

[18] Prior to the change in the propensity to save, the rate of growth of capital and output was the exogenous rate of growth of the labour force. During the 'traverse' from the original balanced growth path to the new balanced growth path, both output and capital per worker are increasing. Thus, during the traverse the rates of growth of capital and output must be *greater* than the rate of growth of the labour force but, as we have seen, in the long run they must revert to the growth rate n. Meade (178) especially p. 42 provides an illuminating alternative exposition of this fundamental point.

Balanced growth in the neoclassical model implies that

$$sf(k) = nk$$

or, dividing both sides of the equation by s,

$$f(k) = \frac{n}{s}k$$

Thus, in Fig. 4.3, balanced growth occurs with the capital–labour ratio k^* which is associated with the intersection of the line nk with $sf(k)$ or, equivalently, with the intersection of the per-worker production function, $f(k)$ with a line drawn with a slope of n/s. (Point D in the diagram.)

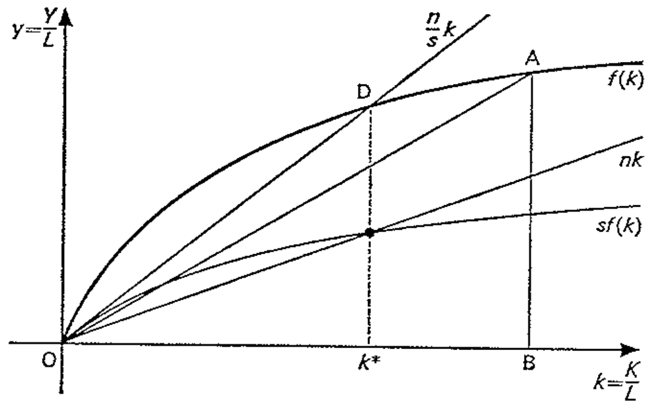

Fig. 4.3

Now consider the slope of *any* line, say OA, from the origin to the per-worker production function. The slope of the line OA is given by

$$\text{Slope of OA} = \frac{\text{AB}}{\text{OB}} = \frac{y}{k} = \frac{Y}{L} \times \frac{L}{K} = \frac{Y}{K}$$

But Y/K is the inverse of the capital–output ratio—i.e. the slope of any line from the origin which intersects the per-worker production function *must* equal $1/v$ given the capital–labour ratio implied by the intersection. Now, we have drawn the line OD with a slope of n/s and we now know that its slope must also equal $1/v$ given the capital–labour ratio k^*. The capital–labour ratio k^* and the associated output per worker therefore together imply that

$$\frac{n}{s} = \frac{1}{v}$$

or, rearranging, that

$$\frac{s}{v} = n$$

which means that the First Harrod Problem is resolved at this point. To each point on the per worker production function there corresponds a different value of the capital–output ratio and the adjustment mechanism of Proposition 1 ensures that the appropriate capital–labour ratio, and therefore the appropriate capital–output ratio, is eventually attained.

In Chapter 3, however, we argued that the constancy of the capital–output ratio in the Harrod model does not arise from purely technological considerations but from the necessity for the rate of interest or the rate of profit to adjust smoothly to produce a varying capital–output ratio. This is illustrated in Fig. 4.4 which, once again, is not as complicated as it may at first appear! In Fig. 4.4, the balanced growth path implies a capital–labour ratio of k^* (i.e. where $sf(k) = nk$ at point H or where $f(k) = (n/s)k$ at point G). Suppose that the actual capital–labour ratio is k^-. We know that, at this point, savings per worker, $sf(k)$, are greater than nk and there is therefore a tendency for the capital–labour ratio to rise. But what behaviour of the wage rate and rate of profit is required by the large change in the capital–labour ratio from k^- to k^*? As shown in 2.4(b)(iv), the rate of profit at k^- is, on marginal productivity assumptions, given by the slope of the tangent AB. To achieve the balanced growth capital–labour ratio, k^*, it must fall (see the slope of the tangent EF) quite considerably as capital becomes less scarce relative to labour. Moreover, the capital–labour ratio does not 'jump' instantaneously from k^- to k^*—it grows gradually and this implies that the rate of profit must fall smoothly as the tangent AB 'moves' around the curve $f(k)$ until it coincides with EF. The factor prices, the real wage rate and the real rental rate on capital, must adjust instantaneously so as to clear the market. It is not necessary to assume the famous Keynesian 'difficulties' of rigid wages or liquidity trap to question whether sufficient information and perception would be available within the markets for capital and labour services to produce the smooth, and possibly very large, transition that would be required during the 'traverse' to the balanced growth configuration of output per worker, y^*, and capital per worker, k^*. Section (v) of Solow's paper investigates the behaviour of interest and wage rates required to sustain the optimistic vision of smooth convergence to balanced growth. He notes that his approach 'directly contradicts Harrod's position that a perpetually falling rate of interest would be needed to maintain equilibrium' (Solow (244) p. 83) and, indeed, it is clear that the fall in the rate of profit (and, under the assumption of the absence or risk, in the rate of interest—see Ch. 3) required is clearly bounded. On the other hand, it is not clear *how long* the adjustment process might take (see 4.7 (ii)) so that the change from k^- to k^* in Fig. 4.4 might require a long period of steadily falling rates of profit and interest before the new steady state is achieved. As Solow admits

(Section (vii) of (244)) the introduction of rigid wages or a liquidity trap could, of course, prevent the necessary adjustments. Rigid wages would allow for the emergence of unemployment in a model in which this is otherwise impossible[19].

(b) The second Harrod problem is completely by-passed by the neoclassical approach. Harrod's instability problem arises from the interaction of his investment function with an implied expectations generating mechanism—allowing cumulative divergences between saving and investment plans. The neoclassical model simply cannot suffer from information-distorting breakdowns of this kind. There is no investment function and no

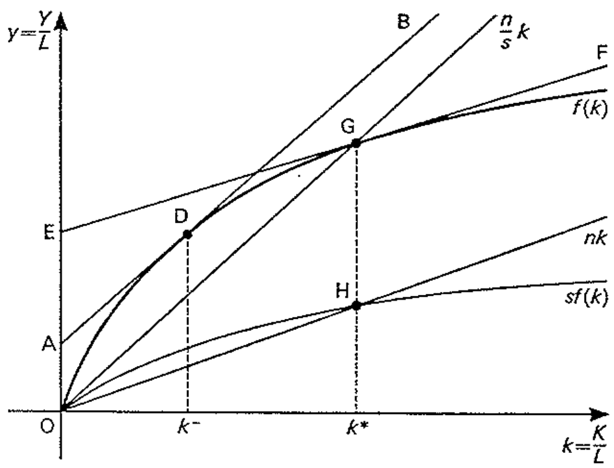

Fig. 4.4

role for the expectations of entrepreneurs. Markets work perfectly and instantaneously. In his original paper Solow suggested that 'if this assumption (that production takes place under conditions of fixed proportions) is abandoned, the knife-edge notion of unstable balance seems to go with it' ((244) p. 65). It is clear that this statement is very misleading. In the words of Stiglitz and Uzawa:

> 'The reason that the Solow model is stable and the Harrod unstable is not because Solow allows for the possibility of substitution of capital for labour, but because of different assumptions about dynamic adjustment and the determination of aggregate output . . . changes in output in the Harrod model depend upon particular assumptions about entrepreneurial behaviour and expectation formation' (Stiglitz and Uzawa (258) p. 13).

[19] In deriving the fundamental equation, the supply of labour is implicitly assumed equal to the demand for labour. Given the neoclassical assumption of perfectly flexible wages, unemployment cannot emerge in this model. But see Conlisk (48).

The Harrod instability problem disappears in a neoclassical model because of the absence of an independent investment function arising from the assumption that ex-ante saving simply *is* ex-ante investment, or that they are maintained in perpetual equality by the monetary and fiscal policies of an omnipotent and omniscient government. As Sen has commented:

> 'Once an independent investment function is introduced, the instability problem of Harrod quickly reappears in the Solow–Swan model, *in spite* of replacing the assumption of a constant capital–output ratio by a neoclassical production function. (See Eisner (63); Hahn (84) and Sen (236)) (Sen (237) p. 23 emphasis added and references altered.)

We can summarize the relationships between the Harrod problems and the simple neoclassical model discussed in this chapter:

(1) *The first Harrod problem is removed by the assumption of a neoclassical aggregate production function implying a variable capital–output ratio, v, together with the assumption of perfect factor markets.*

(2) *The second Harrod problem is by-passed as a result of the absence, in the neoclassical model, of an independent investment function such that the expectations of entrepreneurs have no influence on the economy in general and on the determination of aggregate demand in particular.*

4.6 Extensions of the Simple Neoclassical Model of Economic Growth

The basic neoclassical model of a growing economy can be extended in a number of directions and the assumptions altered in a variety of ways. We list most of these possibilities, elaborate on some, and, in passing, note the role of some of the assumptions utilized.

(a) Depreciating Capital Stock

The fundamental equation derived in Section 4.3 can be easily modified to incorporate the effect of a depreciating capital stock. We redefine I to refer to *gross* investment and assume that the capital stock depreciates at a constant rate λ.

We can therefore write

$$I = \dot{K} + \lambda K \tag{4.6.1}$$

i.e. gross investment equals net investment ($\dot{K}$, the rate of change of the capital stock) plus depreciation. Dividing by L we obtain

$$\frac{I}{L} = \frac{\dot{K}}{L} + \lambda\frac{K}{L} \tag{4.6.2}$$

From equation (4.3.7), we know that $\dot{K}/L = \dot{k} + nk$. Substituting into equation (4.6.2) we obtain

$$I/L = \dot{k} + nk + \lambda k$$

or,

$$I/L = \dot{k} + (n + \lambda)k$$

from which, following the procedure of sections (c) and (d) of 4.3, the fundamental equation is seen to be

$$\dot{k} = sf(k) - (n + \lambda)k \tag{4.6.3}$$

Thus, all the analysis of Section 4.4 can be carried out by replacing n by $(n + \lambda)$.

(b) Neoclassical Growth with a Cobb–Douglas Production Function

The Cobb–Douglas production function is 'well-behaved' and all the preceding analysis could therefore be duplicated using the Cobb–Douglas form of the aggregate production function rather than the general form on which our exposition has been based. Much of Meade's analysis ((178) Chs 2–4) is premised on the implicit assumption of a Cobb–Douglas production function and the interested non-mathematical reader is advised to consult his work[20]. The Cobb–Douglas form of the aggregate production function is particularly tractable mathematically because its multiplicative form implies that logarithmic differentiation to obtain rates of growth is a simple and natural procedure. A detailed study of neoclassical growth with a Cobb–Douglas aggregate production function is available in Hamberg (91) pp. 42–55.

(c) Technological Progress

The neoclassical model of economic growth can easily be extended to include the effects of a simple representation of technological progress. This is discussed in Section 7.4(ii).

(d) Neoclassical Growth and Money

The neoclassical approach can be adapted so as to capture some of the explicit effects of the introduction of money into a growth model. Much

[20] Proposition 1 of Section 7.4 includes an heuristic discussion which, in essence, duplicates much of the formal analysis that would be required in a Cobb–Douglas version of neoclassical economic growth.

of this analysis is due to Tobin (268) but an excellent introductory exposition, utilizing the same tools and similar notation as Sections 4.3 and 4.4, is available in Johnson (116). It could, however, be argued that neoclassical models are fundamentally concerned with the long run and that monetary influences should be ignored in such a context. However, some particularly interesting models have been developed in the advanced literature. (See, for example, Foley and Sidrauski (77) and the references therein.)

(e) Differential Savings

It is possible to analyse a neoclassical model which includes a differential savings function (see 2.3) of the form

$$S = s_w W + s_p P$$

Such an analysis is sketched in Allen (7). We discuss differential savings in the context of the 'Cambridge Criticisms' (see 6.3).

(f) Variations on the Neoclassical Theme

The central neoclassical diagram is particularly easy to modify to demonstrate a number of different propositions:

CASE 1 *Aggregate Production Functions which are not 'well-behaved'*

Figure 4.5 illustrates two per-worker production functions which do not satisfy the conditions of well-behavedness specified in Assumption 4.2.5. Balanced growth requires (see the fundamental equation) that $f(k) = (n/s)k$. $f^1(k)$ is a production function which only intersects $(n/s)k$ at the origin so that no economically meaningful balanced growth path exists. $f^2(k)$

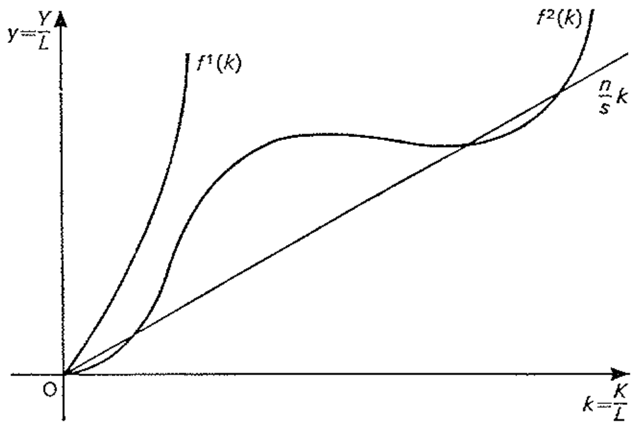

Fig. 4.5

illustrates the case in which multiple balanced growth solutions are possible. The *power* of the assumption of well-behavedness should now be clear. The 'Inada conditions', by constraining the per-worker production function to the shape of Fig. 4.1, ensure[21] that the line $(n/s)k$ has a unique intersection with $f(k)$ (ignoring the origin)—and a unique balanced growth path exists.

CASE 2 *The propensity to save as a function of capital and income per worker*

The simple neoclassical model assumes that the propensity to save, s, is a constant. Figure 4.6 illustrates an interesting situation in which the savings ratio differs at different levels of capital and output per worker[22]. At low levels of capital per worker the savings ratio is low. At medium levels, it increases dramatically whereas, when a high level of capital and output per worker is attained, it 'levels' off again. Consequently, the curve of savings

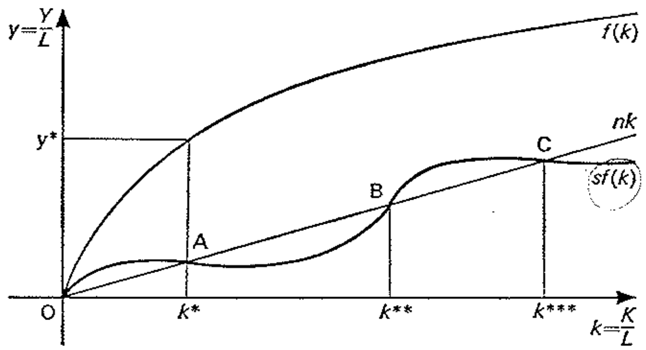

Fig. 4.6

per worker, $sf(k)$, has the shape illustrated in Fig. 4.6 although the per-worker production function is of a well-behaved form. The $sf(k)$ curve intersects the line nk at three points: A, B, and C corresponding to the three levels of the capital–labour ratio, k^*, k^{**}, and k^{***}. Consideration of the fundamental equation

$$\dot{k} = sf(k) - nk$$

shows that the intersection A implies a *stable* balanced growth path with the relatively low levels of income, y^*, and capital, k^*, per worker. In the

[21] A rigorous proof of the existence and uniqueness propositions with a 'well-behaved' production function is available in Burmeister and Dobell (34) pp. 25–9 or Wan (274) pp. 37–9.

[22] This example, together with discussion of a model in which the rate of growth of the labour force varies with the level of output per worker, is included in Johnson (116).

range k^* to k^{**} (i.e. between points A and B) nk is greater than $sf(k)$ and the capital–labour ratio must be *falling* back to k^*. This situation is often referred to in the development literature as a *'Low Level Equilibrium Trap'*[23]. A 'Big Push' or 'critical minimum effort' is required if this economy is to achieve high levels of output per worker. If the capital–labour ratio gets above k^{**} then the forces discussed in Proposition 1 of 4.4 will drive the capital–labour ratio to k^{***} and the associated high steady-state level of output per worker. (Between points B and C, $sf(k)$ exceeds nk and the capital–labour ratio is therefore rising.) A 'gift' of capital sufficient to increase the capital–labour ratio to marginally *above* k^{**} will set this economy on its way to high levels of output per worker. But a 'gift' of capital increasing the capital–labour ratio to marginally *below* k^{**} will be insufficient to produce any major change in the prosperity of the community because savings per worker will simply not be adequate to *maintain* the capital–labour ratio at the new level given the growing labour force (see 9.2).

4.7 Objections to the Neoclassical Model of Economic Growth

The 'neoclassical' model presented in Sections 4.2–4 is not, of course, an adequate representation of the most sophisticated and mathematically advanced neoclassical theories of growth[24]. It does, however, exemplify most of the *central features* of the neoclassical 'vision' of the process of economic growth which, as Section 4.1 emphasized, has become, for whatever reasons, the conventional wisdom of the greater part of the economics profession. The components of the 'vision'—continuous aggregate production functions, perfect markets, instantaneous adjustment of factor prices—together with its central proposition—smooth adjustment to a path of balanced growth which is independent of the propensity to save—have constituted the central subject matter of this chapter and should, by now, be fairly clear. But the 'vision' (and the specific models) has not lacked critics, and the criticisms and objections have mounted with the consolidation of the neoclassical approach. We can summarize these criticisms which, to some extent, overlap with one another, in a series of questions and statements:

(i) *Are not the conceptions of an aggregate production function and aggregate 'capital' subject to such fundamental theoretical weaknesses as to vitiate the conclusions of any analysis based on their use?*

The modern 'Cambridge School', and particularly Joan Robinson, would answer this question with a hearty and resounding affirmative. We will discuss their central criticisms in Chapter 6. It is, however, important

[23] Low level equilibrium 'traps' can exist for a variety of reasons. See, for example, Nelson (188).

[24] The one-good model can be presented far more rigorously and much more complicated models can be constructed which retain the neoclassical flavour.

to note at this stage the peculiar properties of 'capital' which are implied by the neoclassical parable. We have already recognized that the one-good parable implies that the capital goods in our model are technically and economically *homogeneous*. Moreover, the capital goods in the neoclassical model are *non-specific* in the sense that they can be combined with any number of workers to perform any particular task. The justification of Proposition 1 of Section 4.4 requires that, given any initial capital–labour ratio at time zero, $k(0)$, the economy can evolve to a new balanced growth capital–labour ratio by some finite time T, $k^*(T)$. There is no recognition in the 'parable' that the 'machines' existing at time 0 had been, at some time in the past, designed to suit a particular capital–labour ratio and must now be adapted to suit a completely different capital–labour ratio. In short, the neoclassical parable implies a conception of MALLEABLE capital which can be transferred costlessly and instantaneously from operation at one level of the capital–labour ratio to any other. In particular, since 'capital' can be 'moulded' from one form to another, the expectations of entrepreneurs need never be incorrect—for they can always costlessly and instantaneously change a 'mistaken' investment in capital into the appropriate capital good. In Harcourt's words, the malleability of capital implies that: 'both specificity and heterogeneity—two essential characteristics of capital goods—may be abstracted from, and the implications of disappointed expectations . . . may be avoided' (Harcourt (94) p. 5). Meade, in his distinguished exposition of a neoclassical model of economic growth, is particularly explicit: 'The assumption of perfect malleability of machinery' implies that 'a certain tonnage of steel which had been constructed into a machine of a given sort . . . could at a moment's notice and without cost be remoulded into another form of machine' (Meade (178) p. 6). In order to highlight the power of the malleability assumption, neoclassical 'capital' is often referred to as 'butter', 'putty', 'toffee', 'ectoplasm', 'leets' or even 'Meccano sets'[25]—all of which terms are intended to graphically capture the almost magical transmutation properties assumed. The assumption of malleability is directly related to the degree of substitutability of capital for labour—a 'machine' designed to be operated by two men can be instantaneously transferred into a machine suitable for operation by any other number. Now, neoclassical economists do not, of course, really believe that real capital goods are malleable in the sense described above. Malleability is, to them, no more than a possibly fruitful assumption employed in the construction of simple and illuminating models of the process of growth. It is, however, important to realize that,

[25] Mrs Robinson introduced the term 'leets' as a satire on Meade's 'steel' (See Robinson (211)). Swan introduced the idea of 'Meccano' sets as a 'scarecrow to keep off the index number birds and Joan Robinson herself' ((262) p. 343). Harcourt has expressed astonishment that any parent could conceive of a 'Meccano' model being transformed costlessly and instantaneously into one of the other models contained in the instruction book!

while it may be reasonable to assume some possibility for substituting labour for capital in production in the long run, the neoclassical malleability assumption would appear to by-pass some of the central difficulties of a real growing economy in which mistakes are made, expectations are unrealized and the *specific historically-given capital stock* is a crucial constraint on the short and medium-term growth possibilities.

(ii) *How long is the 'long run'?*

Propositions 1 and 2 of Section 4.2 together imply that the long-run rate of growth of the neoclassical economy is Harrod's natural rate of growth which is independent of the propensity to save. But how long is this long run? How long might the economy take to move from capital–labour ratio k^- to the balanced growth capital–labour ratio k^* in Fig. 4.1? How long before the economy returned to the balanced growth path following the change in the propensity to save illustrated in Fig. 4.4? As Atkinson has commented: 'While in many cases we know how the major variables change over time, in very few cases do we know how *quickly* they will change' ((15) p. 137). R. Sato ((231) (232)) has tried to investigate this problem by assuming 'plausible' values of the parameters in a neoclassical model and trying to calculate the length of time that it might take for the economy to attain the balanced growth path following a change in the propensity to save induced by fiscal policy. Now, results of this kind clearly depend sensitively on the actual values of the parameters chosen, but Sato has shown that the adjustment time *might* be very long—possibly over a hundred years[26]—following even a relatively modest change in the savings ratio. As Atkinson has commented: 'The speed of convergence makes a great deal of difference to the way in which we think about the model' ((15) p. 137). A policy of increasing the savings ratio might produce an increased rate of growth of capital and of output for a hundred years or more before a new balanced growth path was attained. Moreover, a very long adjustment period could cast further doubt on the usefulness of the concepts of steady state and balanced growth—for one might expect quite long periods of disequilibrium growth. In sum, if Sato's experiments have any general validity, then Propositions 1 and 2 lose much of their practical significance. On the other hand, Conlisk (48) has attempted to show 'that extending the simplest (neoclassical) model to include the possibility of unemployment reduces the adjustment times to an estimated third of their former values' ((48) p. 562). He introduces some short-run Keynesian mechanisms to produce the possibility of unemployment. (See footnote 19 of this chapter). It might be possible to question whether his model remains truly neoclassical in spirit and it is clear that the introduction of some other Keynesian 'difficulties' could completely change the conclusions of his analysis.

[26] For actual results see Sato (232) pp. 384 and 385.

(iii) *The neoclassical parable neglects or by-passes all the central Keynesian insights—in particular the emphasis on the expectations of entrepreneurs as a decisive force in the macroeconomy.*

It is clear that the neoclassical vision of a smooth long-run equilibrating process is substantially at variance with Keynes' short-run vision of the possibility of the permanent unemployment state. The neoclassical growth model, as with other specific manifestations of the neoclassical spirit, provides what *can* be interpreted as a firm intellectual foundation for discouraging attempts by governments to influence economic processes and problems[27]. As we have already stressed, the 'parable' of economic growth by-passes any possibility of discrepancy between ex-ante savings and ex-ante investment and an independent investment function is not included. In Sen's words:

> 'The assumption of substitutability does not seem to be a key difference between neoclassical and neo-Keynesian studies of growth, though it is sometimes thought to be so, and the main difference seems to lie in the investment function' ((237) p. 23).

In the absence of an investment function, investors' expectations play no role in the determining the growth of the economy. Moreover, the assumption of malleable capital implies that expectations could never be disappointed. Markets are perfect and the market prices of goods and factors accurately transmit the correct information as to relative scarcities. The relative prices of capital and labour change instantaneously to the new *equilibrium* prices following a change in any constant in the model[28].

A dedicated neoclassical economist might argue that ignoring Keynesian 'difficulties' is the appropriate procedure in a long-run model. Many writers, however, feel unhappy with the kind of sharp distinction that is drawn between the difficulties and problems of the short and medium term and the asymptotic optimism of the long run. On the other hand, many neoclassical economists are perfectly aware of the dangers of concentrating on long-run steady states. Samuelson and Modigliani have commented:

> 'We emphasize the centuries that may be involved to stress that we are talking here and everywhere of hypothetical steady states which will never quite be reached from other states and which may be closely approximated only after such long periods of time as to make the models' realism questionable' ((224) pp. 286–7).

but, as they emphasize, this criticism applies to almost *all* modern models of economic growth.

4.8 Conclusion

The simple neoclassical model studied in this chapter provides an apparently consistent and coherent view of the world. The central propositions of the model reverse or by-pass what are described as the 'dismal', 'gloomy' or 'masochistic' prophecies of the Harrod–Domar line of thought (see Eisner (63) p. 707). The model's implications are consistent with the general neoclassical vision of an harmonious, self-equilibrating and mechanistic approach to economic processes.

The neoclassical achievement in the area of growth theory is undeniable and, as we noted at the beginning of this chapter, its influence pervades the way economists think about economic growth. It is, however, important to reiterate that we have studied one of the simplest neoclassical parables of economic growth. Some, perhaps many, would defend it as an illuminating approximation of the reality of a growing economy but it is increasingly generally accepted, not least by many writers who would conventionally be described as 'neoclassical', that the omission of a role for entrepreneurial expectations *is* serious. Commenting on Solow's 1956 assertion that 'when the results of a theory seem to flow specifically from a special crucial assumption that is dubious, the results are suspect' ((224) p. 65), Stiglitz has written:

> 'It now appears—in the perspective of some fifteen years of subsequent research—that the theory developed by Solow, the picture of an economy smoothly converging to balanced growth in an economy in which expectations play no explicit part, is as suspect in this respect as the earlier theory of Harrod' ((260) p. 160).

[27] This is not to imply that writers who have constructed neoclassical models of economic growth would necessarily subscribe to all the prescriptions of the more thorough-going neoclassical economists. But, ideas of a 'natural' rate of unemployment, or a constant rate of growth of the money supply combined with a minimum of government intervention, can be said to share the same vision as the neoclassical growth model.

[28] Eisner (63) was one of the earliest critics to specifically attack the pre-Keynesian aspects of the neoclassical growth model. The 'new' view of Keynes (see (45) or (158)) would emphasize the difficulties of markets working in the neoclassical manner in the absence of a Walrasian 'auctioneer'—the 'deus ex machina' who instantaneously finds the equilibrium prices in a disequilibrium situation.

Comincia prima di cap. 6.

participated actively in the recent debates. Their allies have included a contingent of Italian economists—Garegnani, Nuti, Spaventa—such that Samuelson, in the latest edition of his famous textbook, prefers the description 'Italo-Cantabridgian School' although he is anxious to point out that it 'has able supporters all over the world, including India' ((227) p. 852). Professors Samuelson and Solow, both of M.I.T., have been the principal defenders of neoclassical economics[4] against the attacks of the Cambridge School and they have also been supported by a multinational team of colleagues. It is difficult to avoid gladiatorial imagery when discussing the two camps including, as they do, some of the leading economic theorists of the past thirty years. It is as though two armies have been facing each other for twenty years in a war partly of attrition and partly of swift attack and counterattack.

Two principal intellectual influences constitute the underlying foundation of the Cambridge School's approach to these debates:

(a) Keynes and the 'Keynesian Revolution'

The late Lord Keynes is, in a very real sense, the father-figure of the modern Cambridge School. Many of the most prominent figures were members of the famous 'Cambridge Circus' of young disciples of Keynes who subjected his theories to sustained appraisal and criticism prior to the publication of the *General Theory*—the introduction of which includes a specific acknowledgement of the assistance rendered by Kahn and Robinson[5]. The influence of Keynes' whole method of thinking, together with specific concepts and tools of analysis attributable to him, on the work of the Cambridge School cannot be underestimated. Some writers have suggested that their apparent liking for fierce controversy, together with the style of their denunciation of doctrines which they view as misleading and dangerous, stems directly from their close association with him during the debates over the General Theory when 'Keynes' views . . . were . . . regarded as the ravings of a maniac' (Robinson (210) p. 91) by a large part of the economics profession. Whether their Keynesian inheritance should be regarded as valuable or not is a matter of opinion but it is clear that they have never accepted the gradual emasculation of Keynes' vision to conform with the neoclassical method of thinking—although Mrs Robinson has, in apparent agreement with the so-called 'neoclassical synthesis', argued that: 'as soon as Keynes' views . . . become orthodox doctrine, a large part of them ceased to be relevant.' ((210) p. 91). She is, nevertheless, one of the most fierce critics of what she calls 'Pre-Keynesian theory after Keynes'. The

(4) It is, however, important to note that they would disagree with the Cambridge School as to exactly what it is that they are defending. This should become clear as the chapter proceeds.

(5) For an account of the activities of the 'Circus', together with copies of the voluminous correspondence between Keynes and some of its members, see (142).

specifically Keynesian[6] content of the work of the Cambridge School should become clear as the chapter proceeds.

(b) Ricardo and Marx

A second strand of thought contributing to the approach of the Cambridge School, and one which has become more prominent in the recent debates, stems from the work of the great 'classical' economists—Ricardo and Marx. Sraffa spent many years editing Ricardo's papers into a single collected edition (255) and it is hardly surprising that his own *Production of Commodities by Means of Commodities* ((256)), from which stemmed some of the most important elements in the 'capital controversy' (see 6.2), follows the Ricardian method of analysis[7]. Mrs Robinson has, until recently, been one of the few non-Marxist economists to take Marx's contribution to economic theory seriously[8] and her own work, in particular *The Accumulation of Capital*, shows constant signs of Marxian concepts and modes of thought. It can, in fact, be argued that many of the modern controversies were accurately anticipated by Marx and can be most fruitfully interpreted in the light of an appreciation of his work[9]. Moreover, the Cambridge School in general, and Mrs Robinson, in particular, are, like the classical economists, more inclined towards what we have described in Chapter 1 as the 'grand' theory in which a variety of sociological, psychological and historical facts and theories are intertwined with economic factors in the construction of a model

Any attempt to adequately summarize the contributions of the Cambridge School and the debates themselves in the confines of a short chapter is fraught with a series of serious difficulties[10]:

(i) The Cambridge School cannot be properly regarded as homogeneous in its views. Although the individuals in the group share a similar vision

(6) Following Leijonhufvud's strictures (158), it has become necessary to distinguish between 'Keynesian' economics—i.e. the mechanical five or six equation model which purports to represent the sum total of his thought and which dominates the textbooks—and Keynes' original 'vision' of how the macroeconomy works. The modern Cambridge School is associated far more with his original ideas than with their later formalization and the description 'Keynesian' in Leijonhufvud's sense, should not really be applied to them.

(7) Although Sraffa has emphasized that it is in 'Quesnay's *Tableau Économique* that is found the original picture of the system of production and consumption as a circular process and it stands in striking contrast to the view presented by modern theory of a one-way avenue that leads from "Factors of Production" to "Consumption Goods" ' (Sraffa (256) p. 93).

(8) See, for example, Robinson (213) and Essays 1 and 2 of (210).

(9) See Bhaduri (22) and, less directly, Morishima (187).

(10) Fortunately, a number of full-length expositions are now available. Harcourt (94) is a comprehensive and, unusual for an economist, witty survey of the capital controversies. It is frequently referred to in this chapter. Kregel (151) and the shorter (152) are two surveys of theories of capital and growth written from what is essentially a Cambridge point of view.

of the economic system and are, in general, irritated by the same features of neoclassical economics, they differ in the emphasis they place on various problems in the theory of capital and growth and they are not averse to criticizing one another as well as the common 'enemy'. Thus, there are few occasions on which it is safe to speak of *the* unambiguous view of the Cambridge School.

(ii) The debate between the two Cambridges has generated a vast and often difficult literature. The controversies have raged over twenty years and the ground on which the battles have been fought has shifted considerably in that time. Thus, we cannot hope to obtain more than a *glimpse* of what is involved although footnotes and the reading list provide an approach to the further reading with which the interested student should supplement this chapter.

(iii) The two Cambridges have rarely agreed as to exactly what the arguments are about—to the extent that Solow was moved to preface one of his papers with the statement: 'I have long abandoned the illusion that participants in this debate actually communicate with one another' (Solow (249) p. 207). The lack of communication stems partly from the fact that the two sides are interested in different questions and partly from the different techniques that they employ in making their respective points. Thus, in some of the controversies, the Cambridge School appears primarily interested in the *comparison of the properties of different steady-state or balanced growth paths* while neoclassical growth theorists would not necessarily accept that such comparisons are either particularly interesting or valuable. In terms of techniques, the Cambridge school, and in particular Mrs Robinson, tend to eschew the use of much mathematics in elaborating their arguments preferring what their detractors refer to as a 'woolly cloak of words'[11] but which their supporters would argue avoids the dangers of possibly spurious precision associated with the use of mathematics.

(iv) Underlying methodological and philosophical differences between the two groups as to what constitutes a 'good' model, a 'sensible' assumption or a 'plausible' conclusion tend to lead the two sides to question and argue about the basic *motivation* that led their opponents to construct the theory that they are presenting. Thus, for example, the Cambridge writers are simply not prepared to accept than an economist can incorporate a marginal productivity theory of factor shares in his model without in some sense 'believing' in its general validity. Similarly, the neoclassical school often appears unprepared to accept that the very method of constructing a theory can be interpreted as an attempt to justify rather than analyse. Thus, many of the differences stem from disagreement as to what is a 'scientific' method. Mirrlees has suggested that: 'If anything explains the heat of the debates in growth theory, it is the difficulty thinkers in the

(11) See Lancaster's review ((154) p. 70) of *The Accumulation of Capital.*

scholastic tradition have in appreciating that, for workers in the scientific tradition, it makes sense to entertain a model and use it without being committed to it; while the scientists cannot imagine why mere models should be the object of passion' ((183) p. xxi). There remains, of course, the danger, which is common to all science and which would be emphasized by the Cambridge writers, that the 'mere models' do not remain the property of those who constructed them and who are fully aware of their limitations but may be mingled with other theories, altered, simplified and eventually put to *uses* for which they were never intended. As Keynes commented: 'Practical men, who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some defunct economist' ((140) p. 383).

Despite these difficulties, it is necessary to attempt to provide some insight into the contributions of the Cambridge School; for an exposition, however elementary, of modern theories of capital and growth which ignored them would be analogous to a description of the Wars of the Roses which made no reference to the Lancastrians or to an account of the American Civil War which omitted to consider the existence of the Confederacy. Space constraints preclude a detailed consideration of the different theories and models associated with each member of the Cambridge School[12]. We attempt to isolate some of the common themes that link the different writers together—in particular the so-called 'capital controversy' and their approach to saving (6.2 and 6.3). A common theme that is not accorded a separate section is their attack on the marginal productivity theory of distribution—for their views on this pervade all the other sections. Kaldor's views on technical progress are discussed in Chapter 8. The approach adopted here—of isolating common themes rather than describing the specific models—does have some merits in terms of clarity, simplicity and brevity, but it is important to realize that it suffers from the important weakness of over-emphasizing the critical rather than the constructive aspects of the Cambridge theorizing. That this is so is already implied by the title of this chapter but the reader who requires a broader view is urged to consult the references. On the other hand, it can be argued that some of the most important contributions of the Cambridge School have stemmed from its critical aspects rather than the particular models constructed by individuals within the group.

6.2 The Capital Controversy

Theories of capital have historically been amongst the most fertile sources of economic controversy. Many aspects of the modern debate, if not the techniques employed in its exposition, would have been as familiar to

(12) It can be argued that their models are particularly unsuitable for a summary treatment. Mrs Robinson, for example, has never appeared satisfied with attempts to encapsulate her model. (See Essay 6 of (210), Vol. III).

Marx, Ricardo, Böhm-Bawerk or Wicksell as to any present-day Cambridge (England or Massachusestts) economist. (See Blaug (25) Chs 4, 7, 12.) The interest of economists in the subject of capital is hardly surprising. Since at least the time of the Industrial Revolution, casual observation would have revealed the central role in production of physical capital goods. Moreover, *value capital*, as a form of property conveying some rights to profit, seems to have been, to a greater or lesser extent, a characteristic of human society since time immemorial. But why have theories of capital been *so controversial*? Solow ((250) pp. 9–13) has suggested that controversy in this area arises from two sources: the difficulty of the subject matter and the ideological overtones. Now, there can be no dispute that capital theory is difficult—many of its seemingly simple concepts transpire, on a thorough examination, to be subtle to the point of ambiguity and the mathematical techniques required for the analysis of a rigorous model, incorporating *n* different kinds of capital goods, have become increasingly sophisticated and challenging. It does seem, however, that the principal historical cause of controversy has been ideological rather than a simple matter of comprehension of the theories. The idea that capital is a 'factor of production' whose marginal productivity determines the reward, profit, accruing to its owners is obviously capable of being interpreted as saying something about the 'usefulness' and 'fairness' of the institutional arrangement known as capitalism. Indeed, the crude normative assertions that profits are the 'just' reward for the owners of capital were concocted on the basis of the marginal productivity theory of distribution espoused in the last quarter of the nineteenth century. (See 2.4(b)(iv) and, in particular, the references to Clark.) That controversies should arise in the course of theorizing on the concept of capital and the function of profit is not surprising; that these debates have been so vigorous and virulent cannot be divorced from the general ideological and specific political implications associated with the theories.

In the context of a textbook on theories of economic growth, the central question associated with capital that must be faced is whether the existence in the real world of *heterogeneous* capital goods inevitably invalidates the conclusions of simple theories or 'parables' which incorporate no more than a conception of a *single*, malleable, capital good. The neoclassical conception of capital[13] employed in the one and two-sector models of Chapters 4 and 5, has two functions. Firstly, when, together with the stock of labour employed (or the flow of labour services), it is plugged into an aggregate production function, it is used to explain the flow of output forthcoming. Secondly, as in the models of Chapters 4 and 5, its marginal productivity 'explains' or 'determines' the rate of profit and profits per labourer. All the Cambridge writers have, to a greater or lesser extent, been

[13] Sometimes called the Clark–Ramsey conception of capital. See Clark (42) and (43) and Ramsey (205). As is noted below, modern neoclassical economists would deny that such a conception is necessary for their theories.

concerned to deny that *any unit* can be found in which heterogeneous capital goods can be aggregated so as to simultaneously satisfy the pair of neoclassical requirements described above. Some of them have been more prepared than others to countenance such a possibility or, for practical purposes, to use some concept of aggregate capital in their models, but they are all deeply suspicious of attempts to use aggregate production functions, incorporating aggregate 'capital', so as to explain the flow of output, relative factor prices and the distribution of income. Sraffa's Ricardian method (256) avoids the necessity for the definition of any quantity of aggregate capital and, in his introduction, he specifically notes that, in his approach, 'the marginal product of a factor would not merely be hard to find—it just would not be there to be found' ((256) p. v). Moreover, his book contained one of the first explicit statements of the means whereby 'neoclassical parables' could be shown to be invalid. Kaldor shares the general Cambridge suspicion of aggregate production functions, aggregate capital and the marginal productivity theory of distribution but he would emphasize that the 'difficulties of measurement' (of capital) should not 'lead us to abandon all attempts at measurement' although 'the possibility of measurement depended on the *purpose* for which the measures were required'[14] (emphasis added). Champernowne believed that, apart from the context of unchanging 'stationary states', 'the production function, like Alice's flamingo, will not keep still but wriggles about and regards us with an injured expression as soon as we attempt to make use of it' ((169) pp. 223–4). He was, however, originally prepared to concede that the prospect of constructing an index number of 'capital', satisfying some of the neoclassical requirements, was attractive—but his own famous attempt at constructing a chain index of capital (40) can now be seen to have been instrumental in revealing the Achilles heel of the simple neoclassical conception. The most persistent and resourceful critic of the neoclassical production function and any Clark–Ramsey conception of aggregate capital has been Joan Robinson and the modern debate can be dated from the publication of her famous 1953/4 paper (208) which, in Harcourt's words, 'made a number of specific complaints about the state of economic theory and the state of some economic theorists' ((94) p. 11). Robinson's paper contains a number of related strands of criticism[15] but the most influential allegation was that the neoclassical conception of capital, measured as a single number, simply could not be logically sustained. Conventional present-value methods of valuation (see

[14] Quotations from p. 305 of Lutz and Hague (ed.) (169). Hague's summary record of the debate at the conference on which the volume is based is an extremely valuable statement of Kaldor's views at that time.

[15] Her views on the meaning of 'equilibrium' and on the inappropriate use of theorems relating to the *comparison* of equilibrium positions instead of the analysis of the *process* of moving from one equilibrium to another can, in retrospect, be seen to be more important than a first reading of her paper would indicate. (See Robinson (208) pp. 84–9 and Harcourt (94) pp. 22–3 and 122–3.)

2.2(b)) seemed quite unacceptable to her for such a procedure would imply that

> 'we have to begin by taking the rate of interest as *given*, whereas the main purpose of the production function is to show how wages and the rate of interest (regarded as the wages of capital) *are determined* by technical conditions and the factor ratio' (Robinson (208) p. 81 emphasis added).

Now, a circularity in argument of the kind she describes does seem somewhat unsatisfactory. As long as there exists only a single composite commodity then it is perfectly valid to use an intensive production function, $f(k)$, as in Fig. 4.1 for example, and assert that, in conditions of perfect competition, its slope, given any capital–labour ratio, will equal, or 'determine', the rate of profit. But, if there are a large variety of capital goods which must be aggregated, 'the curve cannot be constructed and its slope measured *unless* the prices which it is intended to determine are known beforehand' (Harcourt (94) p. 20). Mrs Robinson recognized that 'we cannot abandon the production function without an effort to rescue the element of common sense that has been entangled in it' ((208) p. 83) and, for her purposes, she proposed to measure the quantity of capital in terms of the *labour time* required to produce the different heterogeneous items of capital equipment, with a given rate of interest being used to reflect the different gestation periods of different items. As Harcourt has commented: 'this measure has an intuitive appeal as a measure of capital in its role of *productive agent* in capitalist society' ((94) pp. 21-2) (emphasis added) but it is clear that, given a production function with capital measured in terms of labour time (or JR units as Champernowne (40) calls them) 'the wage rate of labour and the reward per unit of capital will, in general, differ under competition from the partial derivatives of output with respect to the quantities of capital and labour employed' (Champernowne (40) p. 112) i.e. the neoclassical marginal productivity theory could *not* be sustained with a Robinsonian measure of capital.

The response to Robinson's strictures was mixed. Solow commented that: 'We have reason to be grateful to her for her annoyance, for she seems to have written her article the way an oyster makes pearls—out of sheer irritation' ((243) p. 101) and his own paper, in general agreement with Robinson, was directed towards showing the 'very narrow class of cases' in which it was reasonable to use a production function with inputs of labour and 'capital in general'. He later wrote that: 'it takes something more than the usual "willing suspension of disbelief" to talk seriously of the aggregate production function' ((245) p. 312). The general response was a retreat into the world of the parable of the single multi-purpose commodity (see Ch. 4) or such ingenious constructs as Swan's 'Meccano' concept of capital which was, in his own words, 'a scarecrow to keep off

the index-number birds and Joan Robinson herself' ((262) p. 343). It was implied that 'profound truths' (Harcourt (94) p. 122) were derivable from the one-commodity parable—that such an assumption was 'remarkably useful' and that the associated model was 'simple but fruitful' (Solow (244) p. 101). Aggregate production functions continued to proliferate in the literature, sometimes with an acknowledgement of the conceptual problems associated with them—'I would not try to justify what follows by calling on fancy theorems on aggregation and index numbers. Either this kind of aggregate economics appeals or it doesn't' (Solow (245) p. 312).

In 1962, Samuelson, in a paper (221) dedicated to Joan Robinson, set out 'to show how we can sometimes predict exactly how certain quite complicated heterogeneous capital models will behave by treating them *as if* they had come from a simple generating production function' ((221) p. 194). His intention was to show that some of the central propositions of what he called the 'neoclassical fairy tale' ((221) p. 201) of the one-good world were *identical* to those generated by a certain kind of model containing a large variety of specific capital goods[16]. In his famous 1956 paper, Solow had commented that: 'The art of successful theorizing is to make the inevitable simplifying assumptions in such a way that the final results are not very sensitive' ((244) p. 65). If Samuelson's argument was, in general, correct then the extreme simplifying assumption of the one-commodity world would have been triumphantly vindicated. It should be clear that Samuelson's argument bears thorough examination.

Firstly, it is necessary to remind ourselves of some of the neoclassical parables in the context of the comparison of different steady states. Consider Fig. 6.1(a). It consists of the conventional per-worker, or intensive, production function—as in the expository diagrams of Chapter 4. Assume that k^- and k^+ represent the capital–labour ratios associated with two different steady states. One could, for example, think of two islands having identical production functions but differing in savings behaviour such that one rejoices in steady-state growth with high levels of output, capital and consumption per worker whereas the other, although also experiencing steady-state growth, has much lower values of these variables. The neoclassical apparatus, together with the marginal productivity theory, provides straightforward predictions concerning the wage/rate of profit configurations in the two steady states:

(a) The rate of profit in the steady state implied by the low capital–labour ratio, k^-, is equal to the slope of the line AE whereas the rate of profit in the steady state implied by the higher capital–labour ratio, k^+, is equal to the slope of the line BD. (See 2.4(b)(iv).) It is therefore clear that the rate of profit associated with the more capital-intensive steady state is *lower*

[16] It must be pointed out that Samuelson himself has been consistent in his insistence 'that capital theory can be rigorously developed without using any Clark-like concept of aggregate "capital"' ((221) p. 193) and has been much less prepared than Solow to use such a concept.

and is therefore employed. For rates of profit between r_1 and r_2 technique beta is employed while for high rates of profit above r_2 technique alpha is once again the more profitable. Now, to those accustomed to the neoclassical vision, the possibilities implied by Fig. 6.4 must seem very paradoxical. The *same* technique (alpha) of producing the single consumption good is the more profitable of the two techniques at both very low and very high levels of the rate of profit. At rates of profit above r_1 the economy 'switches' from technique alpha to technique beta. At rates of profit above r_2 the economy returns or 'reswitches' to technique alpha which had been discarded as unprofitable at intermediate rates of profit. This phenomenon —called the reswitching or double-switching of techniques[21]—clearly casts severe doubt on the general validity of the neoclassical parables which are based upon the explicit assumption that high rates of profit are associated with techniques involving a relatively low level of the aggregate capital–labour ratio. In the neoclassical context, and in the context of the Samuelson model, the same technique simply could not be the most profitable of all techniques at low and high levels of the rate of profit while other techniques are the most profitable at intermediate levels. The simple example of Fig. 6.4 demonstrates that, in general, this presumption cannot be sustained. Two economies, both experiencing steady-state growth, could be employing *exactly the same* technique of production although in one the rate of profit could be relatively low (and the wage rate relatively high) while in the other the rate of profit would be relatively high (and the wage rate relatively low). The 'snark which economists have been hunting ever since Ricardo' (Hicks (108) p. 153)—of being able to 'specify the direction of the change in technique which will be likely to correspond to a lowering of the rate of profit' ((108) p. 153)—has proved to be a quarry so elusive that it seems that, outside the simple parables of the one-commodity world, the quest will always be unsuccessful. It does not take a great deal of intuition to realize that the 'grand factor price frontier' may no longer have the attractive properties of the Samuelson model if the individual wage frontiers of each technique are of a variety of degrees of curvature. Its slope will no longer measure the aggregate capital–labour ratio of the economy and it can be shown[22] that unambiguous statements about capital intensity are almost impossible to make. Its elasticity will no longer equal the ratio of relative shares and, in general, it can be shown that *the presence of reswitching is sufficient to invalidate almost all the neoclassical parables*. (See Harcourt (94) Ch. 4.) Moreover, it is now recognized[23] that it is possible for the neoclassical parables to be invalidated *without* reswitching.

[21] In the context of a more detailed examination of these problems for the case of more than two techniques, it would be necessary to distinguish between 'reswitching' and a related phenomenon known as 'capital-reversing'. (See Harcourt (94) Ch. 4.)
[22] See, for example, Ferguson (72) pp. 259–65.
[23] Following the example provided by Pasinetti (195) pp. 515–16.

The possibility of the reswitching of techniques was first recognized in the literature by Joan Robinson (209), Champernowne (40) and Sraffa (256). Once it became recognized that, in general, it could not be relegated to the status of an unlikely perversity and that it could occur in the economy as a whole[24], its implications for economic theory, and the theory of economic growth in particular, began to be investigated. Some writers (see (33) pp. 534–8 and (259)) have produced sufficiency theorems which demonstrate exactly in what situations reswitching can and cannot occur. The neoclassical school admit that the possibility of reswitching does seriously weaken the attraction of theorizing in terms of parable but do not concede that the new results make any serious difference to the neoclassical edifice *in its full generality*. Nevertheless, the Cambridge writers clearly believe that they have 'won' the argument and, in their own terms, they are surely correct. It is, however, worth recalling that all these arguments relate to the comparison of steady states. Stiglitz (260) has recently shown that in 'truly dynamic economies'—i.e. outside the somewhat artificial compass of the steady state—the really serious problems associated with heterogeneous capital goods are *not* those of reswitching or, in this context, what he calls the 'recurrence of techniques'. These difficulties, which are well outside the scope of an elementary text, relate to the possible or probable non-uniqueness of momentary equilibria (see 5.2) in the presence of heterogeneous capital goods and particular patterns of expectations[25]. The 'capital controversy' has taken a number of different, if related, forms in the twenty or so years of its modern incarnation. It now appears to be over—as a result of the Cambridge writers' belief that their viewpoint has triumphed and the neoclassical writers' contention that the argument was never about the real issues anyway. It is difficult to be certain what is the correct conclusion to draw from these controversies[26]. The simple model of Chapter 4 certainly seems more suspect although some writers, notably Ferguson, have affirmed that they 'have the faith' ((72) p. xvii) to continue to use such constructs. The capital controversies have certainly illuminated many of the difficulties associated with the reality of heterogeneous capital goods and, in particular, the seemingly unassailable problem of deriving correct and unambiguous predictions of

[24] Levhari's attempt (160), at Samuelson's instigation, to show that reswitching could not occur for the economy as a whole was, ironically, the major initiating force in the gradual realization of its importance. Levhari's theorem was conclusively shown to be false by a number of writers. See the symposium (225).
[25] These problems were originally discussed by Hahn and, in the conclusion of his difficult paper, he noted that 'equilibrium dynamics' are 'less attractive once we admit that there are hoes as well as shovels' ((88) p. 646).
[26] The preceding brief discussion of the capital controversy and of the nature and implications of the reswitching of techniques is not, of course, any more than a glimpse of the various issues discussed in the vast literature that has emerged on this subject in recent years. The interested reader is urged to consult the references.

the relationship between steady-state levels of the capital–labour ratio and steady-state factor rewards. In Samuelson's words:

> 'If all this causes headaches for those nostalgic for the old time parables of neoclassical writing, we must remind ourselves that scholars are not born to live an easy existence. We must respect, and appraise, the facts of life' ((226) p. 583).

6.3 Differential Saving—by Income Group and Social Class

The idea that it may be fruitful to distinguish between the propensities to save of capitalists and workers, or between the propensities to save out of different kinds of income, has, at least implicitly, a long history in economics. Most of the 'classical' economists tended to assume that workers did not save. Kaldor has shown ((127) p. 94 footnote 1) how Keynes drew the same sort of distinction in a famous passage of his *Treatise on Money* (139), and much of Kalecki's work[27] employs this assumption in an explicit form. In recent years, this assumption has become one of the distinguishing characteristics of the Cambridge School. All[28] of the Cambridge writers have employed this kind of assumption but it is principally associated with the works of Kaldor and Pasinetti. We examine their arguments in this section although considerations of space preclude a very detailed exposition.

The Kaldor Model

Kaldor's approach to saving, which was originally developed as a Keynesian 'alternative theory of distribution' (127), has become a central element in all his subsequent models of economic growth (see (128), (129) and (130)). It is a simple matter to set out a version of Kaldor's approach to savings. Consider the following equations:

$$Y \equiv W + P \tag{6.3.1}$$

$$S_w = s_w W \tag{6.3.2}$$

$$S_p = s_p P \tag{6.3.3}$$

(6.3.1) states that income, Y, is identically equal to the two broad categories of wages, W, and profits, P. Equations (6.3.2) and (6.3.3) state that savings out of wages, S_w, and savings out of profits, S_p, are proportional to wages and profits respectively. Thus, s_w and s_p are the constant average (and marginal) propensities to save from wages and profits. s_p is assumed greater than s_w.

[27] See, for example, essays 1 and 3 of Kalecki (126).
[28] See Champernowne (41), Robinson (209) and (212) and Kahn (125).

Total savings, S, is given by

$$S = s_w W + s_p P$$

or, substituting for W from (6.3.1),

$$S = s_w(Y - P) + s_p P$$

and rearranging

$$S = (s_p - s_w) P + s_w Y \tag{6.3.4}$$

Now, dynamic equilibrium requires that $I = S$ which implies that

$$I = (s_p - s_w) P + s_w Y \tag{6.3.5}$$

Equation (6.3.5) can be divided by Y and, after a little rearrangement we obtain

$$\frac{P}{Y} = \frac{1}{s_p - s_w}\frac{I}{Y} - \frac{s_w}{s_p - s_w} \tag{6.3.6}$$

Similarly, dividing equation (6.3.5) by K and rearranging we can obtain

$$\frac{P}{K} = \frac{1}{s_p - s_w}\frac{I}{K} - \frac{s_w}{s_p - s_w}\frac{Y}{K} \tag{6.3.7}$$

Now, P/Y is the share of profits in national income and P/K is the rate of profit. In dynamic equilibrium they are seen to be related to the savings propensities out of wages and profits. In the special 'classical' case, employed in our discussion of the Uzawa two-sector model, the propensity to save out of wage income is assumed to be zero and equation (6.3.7) consequently reduces to

$$\frac{P}{K} = \frac{1}{s_p} \cdot \frac{I}{K} \tag{6.3.8}$$

Now, I/K is the rate of growth of the capital stock and in a situation of steady growth at full employment it would equal the natural rate of growth n. Thus, using the assumption that the natural rate is exogenously given, equation (6.3.8) shows that the rate of profit in a state of balanced growth is 'determined' by the propensity to save from profit income. Much the same result was noted by Kalecki in his early work:

> 'Thus, capitalists, as a whole, determine their own profits by the extent of their investment and personal consumption. In a way they are masters of their own fate' ((126) p. 13).

For Kaldor, equations (6.3.6) and (6.3.7) constitute an alternative to the marginal productivity theory of distributions. His critics have not been so impressed. Some have argued that 'Mr Kaldor's system can be regarded as a special case of the neoclassical "marginal productivity" theory' (Findlay (73) p. 178), while Tobin, in a satirical spirit, derived a '*general* Kaldorian theory of distribution' in which there existed n classes of people including 'Actors, Bird-watchers, Conservative Peers, Dons, Executives, Farmers, Gourmets not elsewhere included, Hoopers . . . Nuclear Physicists' ((267) p. 120).

The Kaldorian savings theory provides the Cambridge writers' escape route from the First Harrod Problem. Recalling the requirement for steady-state growth at full employment in the Harrod model

$$\frac{s}{v} = n$$

what we have called the First Harrod Problem stems from the assumption that s, v, and n are all independently determined constants. The neoclassical model solves the problem by making the capital-output ratio, v, a variable (see 4.5). Kaldor's approach implies that the overall average propensity to save, s, is no longer a constant. Dividing equation (6.3.4) by Y we obtain

$$\frac{S}{Y} = s = (s_p - s_w)\frac{P}{Y} + s_w \tag{6.3.9}$$

and it is clear that, within certain limits[29], there will exist a ratio of profits to income, P/Y, that will ensure that the overall average propensity to save is exactly that required to equate s/v to n. Moreover, Kaldor has repeatedly argued that there are reasons to believe that the appropriate value of P/Y will, in fact, emerge:

> 'the "warranted" and the "natural" rates of growth are not independent of one another; if profit margins are flexible, the former will adjust itself to the latter through a consequential change in P/Y' ((127) p. 97).

Thus, Kaldor's explanation for the 'stylized fact' of approximately steady growth at full employment in Western economies in the post-war period is that the distribution of income has been appropriate. Neoclassical writers are sceptical of this view:

> 'Is it teleological shifts in the distribution of income between thrifty and thriftless that, in some run of time, assures a stylized performance of high

[29] Stemming principally from the fact that P cannot exceed Y. See Kaldor (127) where the limits are thoroughly discussed.

> employment with reasonable price stability? If you can believe (this), you can—as the Duke of Wellington once said—believe anything' (Samuelson and Modigliani (224) p. 294 footnote 3).

Despite these criticisms, Kaldor's approach is undoubtedly attractive in its simplicity and it has been utilized in a variety of contexts.

The Pasinetti Model

Pasinetti's contribution was originally motivated by a desire to correct what he called a 'logical slip' in Kaldor's theory but, in so doing, he developed what is a very general model incorporating a surprising and, perhaps, paradoxical conclusion. Pasinetti pointed out that 'in any type of society, when any individual saves a part of his income, he must also be allowed to own it, otherwise he would not save at all' ((194) p. 270). By working explicitly in terms of a division of society into capitalists and workers, it is clear that some part of total profits must accrue to workers as a result of their past savings. Pasinetti reformulated Kaldor's model so as to reflect this observation and his systems of equations are rather similar to those of Kaldor. There is, however, an advantage in presenting the Pasinetti model in an explicitly neoclassical form—partly because such a presentation highlights the generality of Pasinetti's theorem and partly because most of the steps in the derivation are familiar to us from Chapter 4[30]. The assumptions of the model require little comment as most of them were discussed in the exposition of the neoclassical model in Chapter 4.

ASSUMPTION 6.3.1 TECHNOLOGY

Output is produced by capital and labour and the process can be summarized by a neoclassical, well-behaved, constant returns to scale production function. In the intensive form this is

$$y = f(k) \tag{6.3.10}$$

where $y = Y/L$ and $k = K/L$. We assume, for convenience, that capital does not depreciate and there is no technical progress although both these features can easily be incorporated into the model.

ASSUMPTION 6.3.2 OWNERSHIP OF CAPITAL

There are two distinct classes in society—capitalists and workers. Both groups own capital. Thus, the total capital stock, K, is identically equal to the sum of workers' capital, K_w and capitalists' capital, K_c:

$$K \equiv K_c + K_w \tag{6.3.11}$$

[30] The neoclassical presentation of Pasinetti's theorem was originally derived by Samuelson and Modigliani (224).

chapter 7
Simple Conceptions of Technical Progress

7.1 Introduction: Invention, Innovation and Technical Progress

The models of economic growth discussed in Chapters 3–5 of this book have, despite their apparent differences of substance and style, implied that increases in aggregate output and income stem solely from increases in the quantities of the inputs, capital and labour, employed. A cursory reading of the literature on real economic growth or the slightest acquaintance with the economic history of growing economies would suggest that our models, in their emphasis on the accumulation of inputs, have neglected a central element in the process of economic growth: technological change. In this and the following chapter we attempt to rectify this omission by discussing the meaning of technological change, examining methods whereby it can be represented in aggregate models of growth and investigating the consequences of its inclusion in the models of growth discussed in previous chapters.

A discussion of the general concept of technological change is hampered by the variety of meanings that can be attached to it and associated terms such as 'invention' or 'technical progress'. As Schmookler has commented:

> 'Technological change is the "terra incognita" of modern economics . . . we do not even have an agreed upon set of terms' ((234) p. 3).

In the following chapters we will attempt to conform to a consistent set of definitions which are, in the main, derived from the works of Schmookler (234) and Mansfield (171).

We may define *technology* as the 'social pool of knowledge of the industrial arts' (Schmookler (234) p. 1) and the *rate of technological progress* as the rate at which this stock of knowledge is increasing. It is convenient to distinguish between the *effects of technological change and the technological change itself*[1]. Although technological change has its 'more sombre side' (Mansfield (171) p. 3) we will generally assume that its effect is *technical progress* by which we mean that either

(a) More output can be produced given the same quantities of the inputs (or, equivalently, the same amount of output can be generated by smaller quantities of one or more of the inputs); or

[1] This distinction is employed by Kennedy and Thirlwall in their excellent survey of the literature on technical progress (137).

(b) Existing outputs undergo qualitative improvement; or
(c) Totally new products are produced.

An invention which leads to a new technique for producing an existing good is referred to as a *process invention* whereas an invention which changes the form of existing goods or generates totally new goods is referred to as a *product invention*. Once an invention is actually put into practice in the economy a product or process *innovation* is said to have occurred.

Since the mid 1950s, a vast amount of research has been carried out on both microeconomic and macroeconomic aspects of technical progress. The microeconomic studies have investigated the sources of inventions and their causes, together with the speed at which new ideas and methods are disseminated throughout the economy[2]. We are concerned, however, with the role of technical progress in the aggregative context of a growing macroeconomy. Four[3] central questions arise.

(i) *How important is technological and technical progress in the process of economic growth?*

This is a partly theoretical, partly empirical question. Conceptions of technological progress played very little part in economic theory until comparatively recently although Marx incorporated such ideas as a central feature of his analysis of the laws of motion of capitalism and the falling rate of profit (see Blaug (25) Ch. 7) and Schumpeter's thesis of 'waves of innovation', (see, for example, (235) Ch. VIII) although outside the mainstream of conventional economic theory, proved influential in some quarters. Recent interest has stemmed principally from the empirical work of the 1950s (see Section 7.5) which suggested that technical progress was the *most important* factor in determining the rate of growth of the economy. *Consequently*, we are not only interested in the effects that technical progress *could* have in the context of our theoretical models but also in the ways in which these models can be adapted for the investigation of the actual effect of technical progress in the 'real world'.

Whatever the effects of technological change in the process of economic *growth* most writers see it as a pervading influence in the economy. Schmookler identifies it with 'the growth of man's mastery over nature' and Mansfield begins his well-known textbook (171) by asserting that:

> 'without question, technological change is one of the most important determinants of the shape and evolution of the economy. Technological change has improved working conditions, permitted the reduction of working hours, provided an increased flow of products, old and new, and added many new dimensions to our way of life' ((171) p. 3).

[2] See, for example, Mansfield (171) and (172) together with the references contained therein.

[3] (i), (ii) and (iv) rephrased as statements, constitute the 'stylized facts' of technical progress in Nordhaus' study ((190) pp. 8–13).

(ii) *What is the cause of technical progress—is it exogenous or endogenous to the economic system?*

If technical progress is assigned the central role suggested by (i) above, then it is clearly important that its causes be investigated. Most simple models of economic growth have, however, assumed that technical progress proceeds at an *exogenous* rate. As Nordhaus has commented:

> 'Although in most modern price and growth theory technological change is treated as exogenous this must be interpreted as an analytical convenience rather than as a serious statement about the economic system' ((190) p. 9).

Much of the research into the causes of technical progress has been microeconomically inclined, with particular emphasis being placed upon expenditures on research and development (R and D) by firms and industries. Macroeconomic theories incorporating an endogenous conception of technical progress have been developed and some of these are discussed in the next chapter (see 8.3).

(iii) *How is technological change transmitted into real technical progress in the macroeconomy?*

The simplest representations of technological and technical progress in models of economic growth do not specify any particular transmission mechanism whereby the increasing stock of social knowledge is translated into the kinds of technical progress described in (a), (b) and (c) above. Technical progress, in the well-known simile, falls 'like manna from heaven' (see Section 7.2). More complicated models conceive of technological progress as being *embodied* in new items of capital equipment and a simple version of a model employing this specific transmission mechanism is discussed in Section 8.2.

(iv) *If technical progress can be classified as labour saving, neutral or capital saving is there any systematic bias in an economy towards any particular kind of technical progress and, if so, why?*

Attempts at the classification of technical progress have, on the whole, stemmed from an interest in its effect upon the distribution of income between capital and labour (see Section 7.3). A number of possible classification schemes have been proposed but modern interest was revived by the realization that most forms of technical progress appeared to be inconsistent with the conceptions of steady-growth employed in the analysis of many simple growth models. In the next chapter we briefly discuss some of the reasons why technical progress might be expected to take on a particular form in the long run.

7.2 The Representation of Technical Progress

Given the idea that technical progress may be an important factor in the process of economic growth, it is necessary to find means whereby it can be *represented* in the simple growth models of Chapters 2–5. In a one-good model of growth it is clear that the only possible effect of technological progress is to allow more of the single multi-purpose commodity to be produced given the inputs of capital and labour. To quote Salter:

> 'The common characteristic of all such (technological) advances is that they lead to a new production function which is superior to its predecessor in the sense that less of one or more factors of production is required to produce a given output' ((220) p. 21).

Thus, in terms of our per-worker production function (see 2.4(b)(iii)) technical progress implies that the production function shifts upwards as in Fig. 7.1.

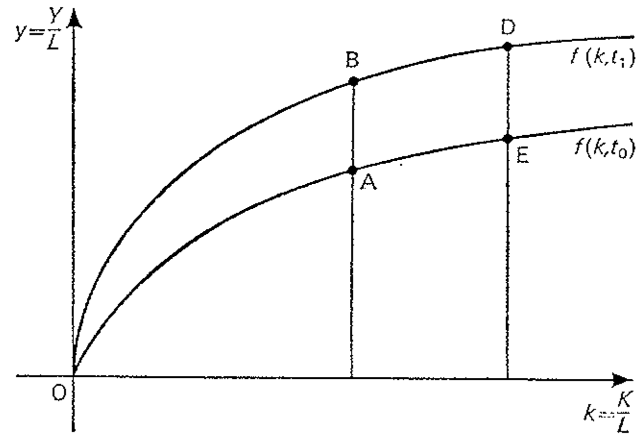

Fig. 7.1

In Fig. 7.1 the production function is originally that illustrated by the curve $f(k, t_0)$. Following technical progress, the curve shifts upwards to the new position $f(k, t_1)$ such that at each level of the capital–labour ratio (except zero) more output per worker can be produced than previously. Although the situation portrayed in Fig. 6.1 constitutes the conventional textbook illustration, Atkinson and Stiglitz (16) have recently emphasized that there is no reason to suppose that the *whole* curve would be shifted upwards by technical progress. They point out that the basic idea underlying the smooth neoclassical per-worker production function is that there exist a large number of different processes of production which can be approximated by a smooth curve (see 6.2). They suggest that technical progress in any one of the separate processes of production need not affect any of the other processes and that, as a consequence, the effect of technical progress would be to produce a 'bulge' in the per-worker production function rather than to shift the whole curve. Their suggestion is illustrated in Fig. 7.2.

The most general method of representing the effect of technological change in a model of economic growth involves rewriting the aggregate production function as

$$Y = F(K, L, t) \tag{7.2.1}$$

Equation (7.2.1) differs from our previous versions of the aggregate production function by the inclusion of the time variable, t, which emphasizes that the output generated by any fixed combination of capital and labour is increasing as time, and technical progress, goes on[4].

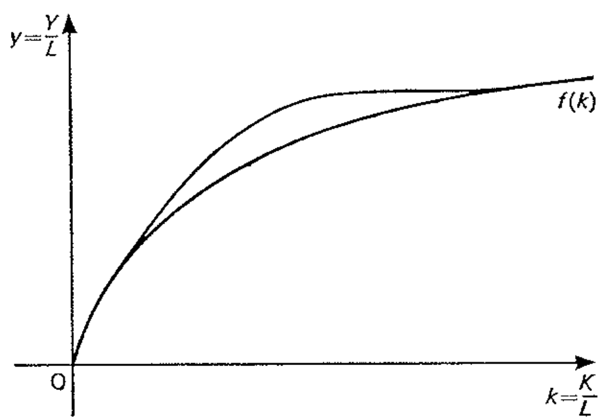

Fig. 7.2

The per-worker production function is written

$$y = f(k, t) \tag{7.2.2}$$

with $\qquad y = Y/L$ and $k = K/L$

Although equations (7.2.1) and (7.2.2) constitute the most general forms of the aggregate production function in the presence of technological progress, a different formulation is widely used in the literature. In this method, technical progress is said to be FACTOR-AUGMENTING. Technical progress shifts the production function such that more output is produced even though the stock of capital and the labour force may not have

[4] The variable t could simply be interpreted as an index of technical progress—but it is convenient to think of it as time.

increased. It is *as if* the factors of production had been augmented. In this formulation the aggregate production function is written as

$$Y = F(A(t)K, B(t)L) \tag{7.2.3}$$

In equation (7.2.3) output, Y, is no longer a simple function of the quantities of capital and labour. The stock of capital, K, and the labour force, L, are multiplied by factors A and B which are both functions of time. The expressions $A(t)K$ and $B(t)L$ are usually referred to as *effective capital* and *effective labour* respectively. The idea is simple. If $\dot{A}(t)$, the rate of change of A, is positive then, as time goes on, the *effective* capital stock increases even though the *actual* capital stock may have remained constant. Similarly, if $\dot{B}(t)$ is positive then the effective labour force is increasing even if the actual labour force is constant. In concrete terms, this form of factor-augmenting technical progress implies, for example, that ten men can do the amount of work previously performed by twelve, and/or five 'machines' produce the output that previously required six.

If $\dot{A}(t)$ is positive and $B(t) = 1$ then technical change is said to be purely *capital-augmenting*. Conversely, if $A(t) = 1$ and $\dot{B}(t)$ is positive then technical change is said to be *purely labour-augmenting*. Finally, if $\dot{A}(t) = \dot{B}(t) > 0$ then technical change is said to be *equally capital and labour augmenting*.

It is crucial to realize that the factor-augmenting representation of technical change does not imply anything about causation or the source of the technical improvement. If, for example, technical progress can be represented as being purely labour augmenting we *cannot* infer that there has been an intrinsic change in the quality of the labour force. In Solow's words: 'It could in fact be an improvement in the design of the typewriter that gives one secretary the strength of 1·04 secretaries after a year has gone by' ((252) p. 35).

Most simple models of growth, incorporating technological progress, assume that technical progress proceeds at a *constant*, *exogenous proportional rate m*. Thus, the factor-augmenting representation of technical progress can be summarized as follows:

$$Y = F(A(t)K, L) \tag{7.2.4}$$

with $\qquad \dot{A}(t)/A(t) = m$

implies purely *capital-augmenting* technical progress at the *constant proportional rate m*.

$$Y = F(K, B(t)L) \tag{7.2.5}$$

with $\qquad \dot{B}(t)/B(t) = m$

implies purely *labour-augmenting* technical progress at the *constant proportional rate m*, and

$$Y = F(A(t)K, B(t)L) \text{ with } \dot{A}(t)/A(t) = \dot{B}(t)/B(t) = m$$

which, by constant returns to scale, can be written as

$$Y = A(t)F(K,L) \tag{7.2.6}$$

implies that technical progress is *equally capital and labour-augmenting* at the *constant proportional rate m*.

The simple procedure of assuming that technical progress proceeds at a *constant proportional* rate is clearly open to severe criticism. Schmookler has commented that: 'Few ideas have proved so intuitively attractive with so little foundation in either logic or evidence' ((234) p. 59). As for the assumption that technical progress is exogenous to the economic system we have already quoted Nordhaus, and Kennedy and Thirlwall are in agreement:

> 'technical progress does not occur by accident but through the deliberate diversion of resources to activities which generate progress in pursuit of fame, profit or both' ((137) p. 13).

Nevertheless, the simple representation of technical progress does provide a starting point for the analysis of models of growth incorporating its influence although its inadequacy as a reflection of 'the real thing' must not be forgotten. We return to these criticisms in Section 7.6.

We can best summarize this section by quoting Hahn and Matthews:

> 'In the simplest treatment technical progress is regarded as *something that goes on at an externally given rate*, and serves to bring about an increase over time in the output that can be produced by any combination of factors of production' ((85) p. 47 emphasis added).

7.3 The Classification of Technical Progress

The interest in classifying technical progress as labour-saving, neutral or capital-saving stems historically from a concern for its effect upon the distribution of income. Two[5] main classificatory schemes, due to Hicks ((107) pp. 121–7) and Harrod ((100) p. 23), exist in the literature and they can both be interpreted in terms of the effect of technical progress upon the relative shares of capital and labour.

[5] A third method of classification due to Solow (247) and Fei and Ranis (69) is briefly discussed below. For more detail see Allen (7) p. 239 or Burmeister and Dobell (34) pp. 68–77.)

(a) Hicks' Classification of Technical Progress

Sir John Hicks introduced the classification of technical progress associated with his name in his *Theory of Wages* (107):

> 'We can classify inventions accordingly as their initial effects are to increase, leave unchanged or diminish the ratio of the marginal product of capital to that of labour. We may call these inventions "labour-saving", "neutral" and "capital-saving" respectively' ((107) p. 121).

In this form the definition appears to be admirably clear and unambiguous. If we denote the marginal products of capital and labour *before* the onset of technical progress as $F_K(0)$ and $F_L(0)$ respectively, and the same marginal products *after* technical progress as $F_K(t)$ and $F_L(t)$ then Hicks' definition can be summarized as follows:

If $\dfrac{F_K(t)}{F_L(t)} > \dfrac{F_K(0)}{F_L(0)}$ then the technical progress is *labour-saving* in the Hicks' classification.

If $\dfrac{F_K(t)}{F_L(t)} = \dfrac{F_K(0)}{F_L(0)}$ then the technical progress is *Hicks-neutral*.

If $\dfrac{F_K(t)}{F_L(t)} < \dfrac{F_K(0)}{F_L(0)}$ then the technical progress is *capital-saving* in the Hicks' classification.

Hicks' classification can be given a simple economic interpretation. We already know that, in competitive conditions, the marginal product of capital is equal to the rental rate on capital and the marginal product of labour is equal to the wage rate (see Section 2.4(b)(iv)). Thus, a labour-saving invention *increases* the ratio r/w (and, obviously, decreases the wage/rental ratio w/r) while a capital-saving invention decreases r/w (and increases w/r). An invention that saves labour implies that the wage rate declines relative to the rental rate of capital which, given that labour is less scarce relative to capital following the invention, is as one would expect.

A crucial problem immediately presents itself. Consider Fig. 7.1. The ratio of the marginal product of capital to that of labour is different at every point on both curves. With the whole per-worker production function having shifted upwards as a result of technical progress it is necessary to *specify which point* on the new curve is to be compared with which point on the old if Hicks' classification is to be of any use. If, for example, the ratio of the marginal products at point D on $f(k,t_1)$ is compared with the same ratio at point A of $f(k,t_0)$ then it is clear that a different classification will be made than if the ratio of marginal products at point B were compared with the ratio at point A.

The Hicks' classification is based upon the comparison of points at which the *capital–labour ratio is constant*, i.e. points A and B (or D and E) in Fig. 7.1. We therefore need to rephrase our definitions of Hicks-neutral, labour-saving and capital-saving technical progress.

DEFINITION 7.3.1

An upward shift, representing technical progress, in the per-worker production function is said to be *labour-saving* (capital-saving), if *at any constant value of the capital–labour ratio*, the ratio of the marginal product of capital to the marginal product of labour has increased (decreased).

DEFINITION 7.3.2

An upward shift, representing technical progress, in the per-worker production function is said to be *Hicks-neutral* if, at any constant value of the capital–labour ratio, the ratio of the marginal product of capital to the marginal product of labour remains constant.

The complete Hicks' classification specified in Definitions 7.3.1 and 7.3.2 can easily be interpreted in terms of the effect of technical progress on the relative shares of national income accruing to capital and labour. We have already noted that Hicks labour-saving (capital-saving) technical progress implies that, in competitive conditions, the ratio of the rental rate on capital to the wage rate, r/w, is increasing (decreasing) and that Hicks-neutral technical progress implies that this ratio remains constant. Now, the ratio of relative shares, which we denote as Π, is equal to rK/wL. For all changes in which the capital–labour ratio, K/L, remains constant it is clear that technical progress, in affecting r/w, will systematically affect the ratio of relative shares. Thus, an alternative way of stating the Hicks classification is simply in terms of the effect of technical progress upon the ratio of relative shares.

Technical progress is said to be *labour-saving* in Hicks' sense if, at any constant value of the capital–labour ratio, *the ratio of relative shares*, $\Pi = rK/wL$, *is increasing*. i.e. $\dot{\Pi}$, the rate of change of relative shares, is positive).

Technical progress is said to be *capital-saving* in Hicks' sense if, at any constant value of the capital–labour ratio, *the ratio of relative shares*, $\Pi = rK/wL$, is decreasing, (i.e. $\dot{\Pi} < 0$).

Technical progress is said to be Hicks-neutral if, at any constant value of the capital–labour ratio, *the ratio of relative shares*, $\Pi = rK/wL$, *remains constant*. (i.e. $\dot{\Pi} = 0$).

A well-known diagram can be used to illustrate Hicks-neutral technical progress. In Fig. 7.3, $f(k,t_0)$ represents a conventional per-worker production function. Assume that the economy has a capital–labour ratio of k^*. Given this capital–labour ratio, the marginal product of capital is measured by the slope of the tangent RA and the wage-rental ratio, $z = w/r$, is measured by the distance OR. (See Section 2.4(b)(iv).) If technical progress causes the per-worker production function to shift upwards to $f(k,t_1)$ then Hicks neutrality requires that, at the capital–labour ratio k^*, the ratio of the marginal product of capital to the marginal product of labour, or the ratio of the rental rate on capital to the wage rate, r/w, must remain constant. Thus, a Hicks-neutral shift from $f(k,t_0)$ to $f(k,t_1)$ requires that the tangent to the new production function at the capital–labour ratio k^* must originate from R such that the distance OR ($= w/r$) remains the same after the shift. These conditions are seen to be satisfied in Fig. 7.3 and the shift from C to D in the production function therefore represents Hicks-neutral technical progress.

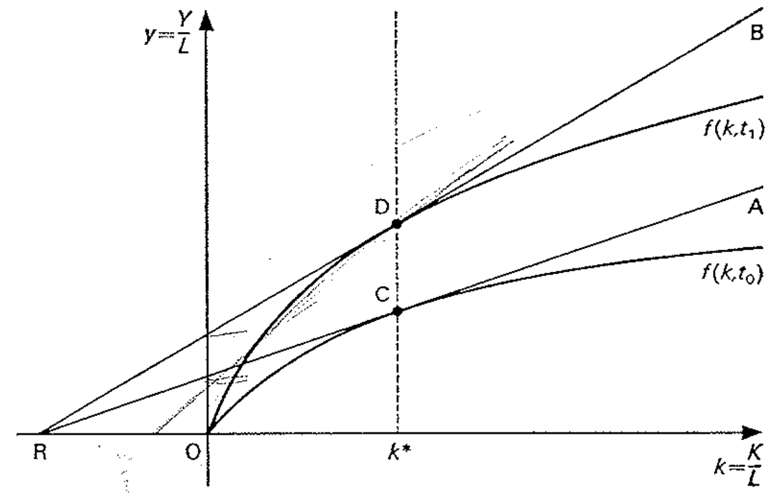

Fig. 7.3

It can be proved (see Uzawa (270)) that Hicks-neutral technical progress is *exactly equivalent* to the idea of equally capital and labour augmenting technical progress discussed in Section 7.2. Thus, if technical progress is proceeding at a constant, Hicks-neutral, proportional rate m then the aggregate production function is identical to that of equation (7.2.6). Alternatively, the concept of continuous growth introduced in Section 2.5 can be employed in which case the aggregate production function takes the form:

$$Y = e^{mt}F(K,L) \tag{7.3.1}$$

Equation (7.3.1) therefore represents an aggregate production function in which technical progress is Hicks-neutral and is proceeding at a constant proportional rate m. We have somewhat laboured the differing interpretations of Hicks' classification of technical change because the equivalent, but seemingly different, definitions employed in the literature frequently confuse the student.

Hicks' classification of technical progress was devised specifically within the context of a theory of wages and distribution—and in that context it can be both useful and illuminating. It is not, however, particularly useful[6] within the kinds of models of economic growth that we have been discussing in previous chapters. Consider the simple neoclassical model of Chapter 4. Figure 7.4 illustrates the result of a shift in the production function caused by technical progress.

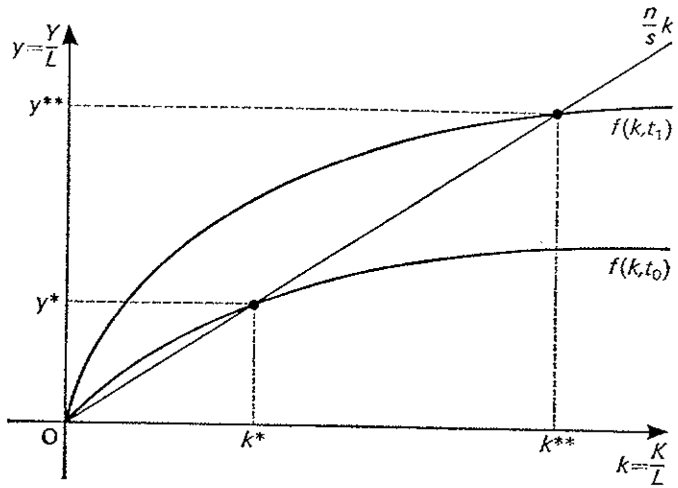

Fig. 7.4

In Fig. 7.4, $f(k,t_0)$ represents the per-worker production function before the onset of technical progress. Given the rate of growth of the labour force, n, and the propensity to save, s, we know (see Proposition 1 of Chapter 4) that the long-run equilibrium of the economy implies a level of output per worker of y^* and a level of capital per worker of k^*, (i.e. where $f(k,t_0) = (n/s)k$). If the production function shifts up to $f(k,t_1)$ as a result of technical progress then the new long-run equilibrium position for the economy is given by the level of capital per worker of k^{**} and output per worker of y^{**}. If the economy was originally on the balanced-growth path implied by k^* then the forces discussed in Chapter 4 are set in motion and the economy gravitates smoothly to the new balanced growth path associated with k^{**}. Fig. 7.4 illustrates a crucial point. If n, the rate of growth of the labour force, and s, the propensity to save, are both fixed constants then *any* upward shift in the per-worker production function will always

[6] And, as we will see in Section 7.4, Hicks-neutral technical progress is not consistent with any conception of steady growth.

imply, in a neoclassical model, a new stable balanced growth path involving a *higher* capital–labour ratio and level of output per worker. Hicks' classification of technical progress is restricted to the comparison of points involving a *constant* capital–labour ratio and will therefore not be useful in the context of conventional steady-state models of economic growth. Harrod's alternative classification of technical progress was devised for use in models of a steadily growing economy and it is to his schema that we now turn.

(b) Harrod's Classification of Technical Progress

Sir Roy Harrod introduced the classification of technical progress associated with his name in a review article in the *Economic Journal* in 1937 (98) but it is best known from his discussion in *Towards a Dynamic Economics* ((100) pp. 22–8):

> 'I define a neutral advance as one which, at a constant rate of interest, does not disturb the value of the capital coefficient' and 'A stream of inventions, which are neutral as defined, will, provided that the rate of interest is unchanged, leave the distribution of the total national product as between labour (in the broadest sense) and capital unchanged' ((100) pp. 23).

The capital coefficient is defined as 'the ratio of the value of capital in use to income per period' ((100) p. 22) i.e. the capital–output ratio, v, and, in competitive conditions and assuming the absence of risk, the rate of interest is equal to the marginal product of capital. Thus, Harrod-neutral technical progress can be defined as follows.

Definition 7.3.3

An upward shift, representing technical progress, in the per-worker production function, is said to be *Harrod-neutral* if, *at any constant value of the capital–output ratio*, the marginal product of capital remins unchanged[7]. Harrod's classification compares points at which the *capital–output ratio* is constant as opposed to the Hicks procedure which compares points at which the capital–labour ratio is constant. A complete statement of the Harrod classification is most easily accomplished by taking up the suggestion included in the second passage quoted above and investigating the effect upon the relative shares in national income accruing to capital and labour as technical progress proceeds.

[7] Notice that Harrod's exact approach is changed in our definition. Instead of determining the neutrality or non-neutrality of an invention 'by reference to what happens to the capital coefficient, if the rate of interest is constant' ((100) p. 27) we have followed the modern procedure of questioning what happens to the rate of interest if the capital–output ratio (the capital coefficient) is constant.

DEFINITION 7.3.4

Technical progress is said to be labour-saving (capital-saving) in Harrod's sense if, at any constant value of the capital–output ratio, the ratio of relative shares, $\Pi = rK/wL$, is increasing (decreasing), i.e. if K/Y is constant and $\dot{\Pi} > 0$ then technical progress is Harrod labour-saving. If K/Y is constant and $\dot{\Pi} < 0$ then technical progress is Harrod capital-saving.

DEFINITION 7.3.5

Technical progress is said to be Harrod-neutral if, at any constant value of the capital–output ratio, the ratio of relative shares, $\Pi = rK/wL$, remains constant, (i.e. $\dot{\Pi} = 0$ when K/Y is constant).

Thus, both the Hicks and Harrod method of classifying technical progress can be reduced to a study of the effect upon the relative shares of capital and labour resulting from the shift in the aggregate production function. But their classifications differ in *which* points on the old and new production functions are compared—Harrod's method being applicable to the comparison of points with equal capital–output ratios while Hicks' system is applicable to the comparison of points with equal capital–labour ratios.

A Harrod-neutral shift in the per-worker production function can be easily illustrated as in Fig. 7.5.

In Fig. 7.5 $f(k,t_0)$ represents a conventional per-worker production function prior to the onset of technical progress. If the economy is initially

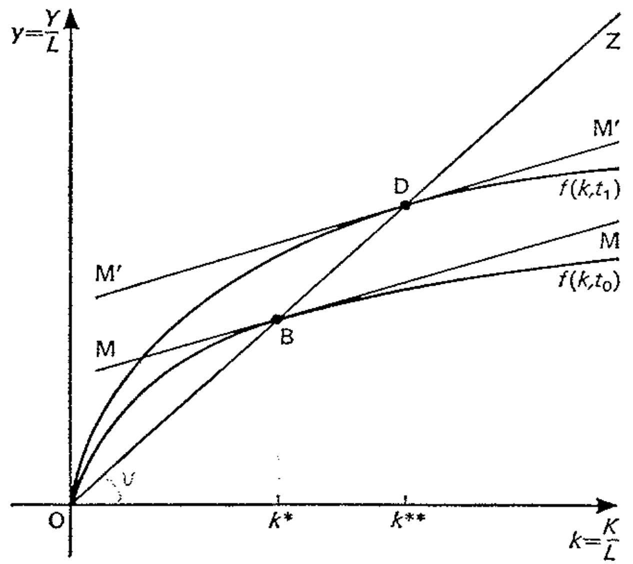

Fig. 7.5

operating at a capital–labour ratio equal to k^* we know (see Section 4.5(a)) that the capital–output ratio is equal to the inverse of the slope of the line OBD (since slope of OBD $= Bk^*/Ok^* = y/k = Y/L \div K/L = Y/K = 1/v$) and the marginal product of capital is equal to the slope of the tangent MM at the point B on $f(k,t_0)$. Technical progress now shifts the per-worker production function to the new position $f(k,t_1)$. If the shift is to be Harrod-neutral two conditions must be satisfied:

(1) the marginal product of capital must remain the same as that given by the slope of the tangent MM and

(2) the capital–output ratio must be equal to that determined by the slope of the line OBZ.

These conditions can be seen to be satisfied in the movement from the point B, with capital–labour ratio of k^*, to the point D, with capital–labour ratio of k^{**}, in Fig. 7.5. At the point D, the capital–output ratio is given by the inverse of the slope of OBZ and is therefore equal to the capital–output ratio at the point B. The marginal product of capital at point D is equal to the slope of the tangent M′M′ and, since this tangent is parallel to the tangent MM at point B, is equal to the marginal product of capital at point B. Thus, the shift from point B to point D in Fig. 7.5 represents Harrod-neutral technical progress. For the shift in the *whole* curve to be generated by Harrod-neutral technical progress it is necessary that the marginal product of capital remains constant *whatever* the constant value of the capital–output ratio (i.e. given the inverse of the slope of any line from the origin through the original and post-technical progress production functions).

In 1938, Mrs Joan Robinson, in response to Harrod's original formulation of his classificatory scheme, demonstrated[8] geometrically that Harrod-neutral technical progress is exactly equivalent to what we have called purely labour-augmenting technical progress (see Section 7.2). In her own words:

> 'a neutral invention in Mr Harrod's sense has the same effect as an increase in the supply of labour . . . and is seen to be equivalent to an all-round increase in the efficiency of labour' ((20) p. 140).

Thus, if technical progress is Harrod-neutral and is proceeding at a constant proportional rate, m, then the aggregate production function may be written as

$$Y = F(K,B(t)L) \tag{7.3.2}$$

[8] We do not prove this proposition. Mrs Robinson's geometrical argument is easy to follow and Allen (7) p. 241 provides a simple demonstration that purely labour-augmenting technical progress implies Harrod-neutrality. A rigorous mathematical proof that if and only if technical progress is Harrod-neutral can the aggregate production function be written in purely labour-augmenting form is provided in Uzawa (270).

with $\dot{B}(t)/B(t) = m$

or, utilizing the concept of continuous growth,

$$Y = F(K, e^{mt}L) \tag{7.3.3}$$

One result of Mrs Robinson's theorem is that Harrod-neutral technical progress, because of its equivalence with an increase in the labour force, is particularly easy to incorporate in the models of growth that we have examined in Chapters 3–6.

It is interesting to note a form of the technology which is consistent with both Hicks and Harrod-neutral technical progress. Consider Figs 7.3 and 7.5. In Fig. 7.3, k^* is the initial capital–labour ratio and Hicks-neutral technical progress maintains the distribution of income constant at this *level* of the capital–labour ratio. In Fig. 7.5, assume that k^* is the initial steady-state capital–labour ratio, (which implies that the line OBZ has a slope equal to n/s—see Section 4.5). Harrod-neutral technical progress implies that the distribution of income remains constant at a constant capital–output ratio which, as can be seen in Fig. 7.5, involves a *higher* steady-state capital–labour ratio of k^{**}. If Hicks and Harrod-neutrality are to be equivalent it is clear that the distribution of income must be the same at every level of the capital–labour ratio. We demonstrate *that this will be true if the elasticity* of substitution (see 2.4(b)(v)) *between capital and labour is equal to unity.* The ratio of relative shares can be written as

$$\Pi = \bar{p}\,.\,k \text{ where } \bar{p} = r/w \text{ and } k = K/L$$

If both $\bar{p}$, the ratio of the rental rate of capital to the wage rate, and k, the capital–labour ratio, are growing then it is clear that Π will be growing. Writing the rates of growth[9] as $\Delta\Pi/\Pi$, $\Delta\bar{p}/\bar{p}$ and $\Delta k/k$ (where Δ means, as in Chapter 2, 'an increment in') then

$$\frac{\Delta\Pi}{\Pi} = \frac{\Delta\bar{p}}{\bar{p}} = \frac{\Delta k}{k}$$

If relative shares are to remain constant $\Delta\Pi/\Pi$ must equal zero and

$$\frac{\Delta k}{k} = -\frac{\Delta\bar{p}}{\bar{p}}$$

[9] A slightly more rigorous analysis would involve taking natural logarithms of $\Pi = \bar{p}\,.\,k$, differentiating the result, setting $d\Pi/\Pi$ equal to zero and, by rearrangement, showing that σ must equal 1.

Cross-multiplying we obtain

$$-\frac{\bar{p}}{k}\cdot\frac{\Delta k}{\Delta\bar{p}} = 1 \tag{7.3.4}$$

The left-hand side of equation (7.3.4) is, of course, the elasticity of substitution (see 2.4(b)(v)) and we have established that, if the ratio of the relative shares in income of capital and labour are to remain constant at any level of the capital–labour ratio, it must equal one[10]. In Chapter 2 we noted that the only form of the technology which has a constant elasticity of substitution of one for every level of the capital–labour ratio is representable by the Cobb–Douglas production function. We have therefore demonstrated that *if the aggregate production function is of the Cobb–Douglas form then technical progress can be interpreted as being either Hicks or Harrod-neutral*[11]. The 'unequivocal neutrality' ((85) p. 51) of the Cobb–Douglas production function has constituted one of the principal reasons for its popularity as a cornerstone of many models of economic growth.

We have examined the Hicks and Harrod classifications of technical progress at some length. We have seen that although both systems can be systematized in terms of the effect upon the distribution of income of technical progress they differ crucially in the kinds of shift in the aggregate production function to which they are applicable[12]. A third system of classification, usually attributed to Solow (247), is identical to that of Harrod and Hicks except that it compares points on the new and old production functions at which the *labour–output ratio* is constant. It can be shown that Solow-neutral technical progress corresponds to what we have called purely capital-augmenting technical progress and that the only circumstances in which it is equivalent to Hicks and Harrod-neutrality is when the technology can be represented by a Cobb–Douglas aggregate production function (see Allen (7) p. 250). Our examination of the different ways of classifying technical progress has been somewhat tedious but it constitutes a necessary preliminary to the attempt at incorporating a simple representation of technical progress in the specific growth models presented in Chapters 3–6.

7.4 Technical Progress in Growth Models

As we have seen, most models of economic growth are concerned, at least

[10] A similar heuristic argument is available in Hahn and Matthews (85) p. 51.

[11] Uzawa (270) includes a rigorous demonstration of these propositions. Allen (7) pp. 248–51 systematically analyses Hicks and Harrod-neutral technical progress at a constant proportional rate of m within the general framework of an aggregate production function of the Cobb–Douglas form.

[12] It is easy to show, using the same kind of reasoning as that used to derive equation (7.3.4) that: (a) if $\sigma > 1$, Harrod-neutral technical progress is equivalent to Hicks labour-saving technical progress and (b) if $\sigma < 1$ Harrod-neutral technical progress is equivalent to Hicks capital-saving technical progress (see Robinson (207) pp. 140–2).

in part, with the possibility of attaining and sustaining a path of steady growth. When we attempt to incorporate the influence of technical progress in these models we are constrained by a simple and perhaps surprising fact: the only kind of technical progress consistent with steady-growth in simple growth models is the special Harrod-neutral or labour-augmenting form.
We state this as Proposition 1.

PROPOSITION 1

If sustained steady-state growth is to be possible in a simple model of economic growth incorporating technical progress, then technical progress must take the Harrod-neutral or labour-augmenting form.

As Solow has commented:

> 'It is not easy to explain why this special labour-augmenting form of technological progress is necessary for steady-state growth to be possible' ((252) p. 35).

A rigorous mathematical proof such as that of Uzawa (270) is not really difficult and the mathematically inclined student is urged to consult his paper. We employ an heuristic argument which should clarify the nature of the problem.

It is necessary, as a preliminary, to demonstrate that, in a general neo-classical model with technical progress, the rate of growth of output per worker is equal to the sum of the rate of growth of capital per worker, weighted by capital's share in national income, with the rate of technical progress.

Consider the most general version of the aggregate production function incorporating technical progress:

$$Y = F(K,L,t) \tag{7.4.1}$$

An increment, ΔY in total output can result from an increment ΔK in the capital stock, an increment, ΔL, in the labour force, the effect of exogenous technical progress encapsulated in the variable t, or any combination of these three factors. If the capital stock increases by an increment, ΔK, then the increment in output so generated will equal the increment in capital multiplied by capital's marginal productivity in producing output. Similarly, an increment in the labour force will generate an increment in output equal to the marginal product of labour multiplied by the additional labour. The total effect of the three factors leading to an increment in output can therefore be written as

$$\Delta Y = \Delta K \times \text{Marginal Product of Capital} + \Delta L \times \text{Marginal Product of Labour} + \Delta Q \tag{7.4.2}$$

where ΔQ represents the increment in output stemming from the effect of technical progress. It is convenient if, instead of specifying discrete increments in the variables, we continue the analysis in terms of the rate of change of each variable. Thus, (7.2.4) can be rewritten as

$$\dot{Y} = \dot{K} . \frac{\partial Y}{\partial K} + \dot{L} . \frac{\partial Y}{\partial L} + \dot{Q} \tag{7.4.3}$$

where $\partial Y/\partial K$ and $\partial Y/\partial L$ represent the marginal products of capital and labour respectively. Equation (7.4.3) can be divided by Y to produce

$$\frac{\dot{Y}}{Y} = \frac{\dot{K}}{Y} \cdot \frac{\partial Y}{\partial K} + \frac{\dot{L}}{Y} \cdot \frac{\partial Y}{\partial L} + \frac{\dot{Q}}{Y}$$

which can be rewritten as

$$\frac{\dot{Y}}{Y} = \frac{\dot{K}}{K} \cdot \left[\frac{K}{Y} \cdot \frac{\partial Y}{\partial K}\right] + \frac{\dot{L}}{L} \cdot \left[\frac{\partial Y}{\partial L} \cdot \frac{L}{Y}\right] + \frac{\dot{Q}}{Y} \tag{7.4.4}$$

Now, in competitive conditions, $\partial Y/\partial K$ and $\partial Y/\partial L$, the marginal products of capital and labour, equal the rental rate of capital and the wage rate respectively. Thus, $\frac{K}{Y} \cdot \frac{\partial Y}{\partial K}$ and $\frac{L}{Y} \cdot \frac{\partial Y}{\partial L}$ equal the shares of capital and labour in national income. If we write the share of capital as Π_K and the share of labour as Π_L (which equals $(1 - \Pi_K)$ since constant returns to scale implies that $\Pi_L + \Pi_K = 1$) and employ the notation $\hat{Y} = \dot{Y}/Y$, $\hat{K} = \dot{K}/K$ and $\hat{L} = \dot{L}/L$ for the rates of growth of output, capital and labour, then (7.4.4) can be written as

$$\hat{Y} = \hat{K} . \Pi_K + \hat{L}\,\Pi_L + \dot{Q}/Y$$

or

$$\hat{Y} = \hat{K} . \Pi_K + (1 - \Pi_K)\hat{L} + \dot{Q}/Y$$

from which we can, with a little rearrangement, obtain

$$\hat{Y} - \hat{L} = \Pi_K(\hat{K} - \hat{L}) + \dot{Q}/Y \tag{7.4.5}$$

In equation (7.4.5) $(\hat{Y} - \hat{L})$ corresponds to the rate of growth of output per worker and $(\hat{K} - \hat{L})$ is the rate of growth of capital per worker. (See the derivation of equation (4.3.5).) $\dot{Q}/Y$ represents the proportional rate of growth of output which is attributable to technical progress rather than any increase in the factors employed. Thus, equation (7.4.5) represents the result that we set out to demonstrate.

Using equation (7.4.5) it is a relatively easy matter to show that any form of non-Harrod-neutral technical progress is incompatible with steady-state growth in this kind of model. Harrod-neutrality requires that, along paths of growth on which the capital–output ratio is constant, the distribution of income between capital and labour must remain constant. Now, if the rate of growth of the capital stock, $\dot{K}/K = sY/K$, is to be constant (steady growth) then, given that the propensity to save, s, is constant, the output–capital ratio, $Y/K = 1/v$, must remain constant which implies that the capital stock is growing at the same rate as output. Thus, a constant capital–output ratio is necessary for steady growth in any simple model of growth. Consider what would happen if technical progress were not Harrod-neutral, in which case either the capital–output ratio is changing or the distribution of income is changing as growth proceeds. If the capital–output ratio is changing then it is clear that the economy is not in a state of steady growth. What if, on the other hand, the capital–output ratio is constant but the distribution of income is changing? The capital–output ratio, v, is equal to

$$v = K/Y = K/L \div Y/L = k/y.$$

Thus, if the capital–output ratio is to *remain* constant the rate of growth of output per worker must equal the rate of growth of capital per worker. Equation (7.4.5) demonstrates that the rate of growth of output per worker is equal to the sum of the exogenous rate of technical progress and the rate of growth of capital per worker weighted by the share of capital. It is clear that, if the share of capital, Π_K, is changing then the rate of growth of output per worker *cannot* remain equal to the rate of growth of capital per worker and steady growth is not maintainable. We have therefore demonstrated that if non-Harrod-neutral technical progress takes place in an economy that was previously experiencing steady-state growth then the steady-state growth cannot be maintained[13].

This result is clearly somewhat worrying for there is no immediately obvious reason why technical progress should be systematically Harrod-neutral in form[14]. If, for example, technical progress is systematically Hicks-neutral and the elasticity of substitution between capital and labour is not equal to one then the maintenance of a constant capital–output ratio (steady growth) implies that the share of either capital or labour tends

[13] The preceding reasoning is based on the discussion by Hahn and Matthews (85) p. 51. Solow ((252) p. 35) employs a neat verbal argument to make the same point.
[14] See, however, Section 8.3(ii).

to zero as time goes on[15]. A variety of responses are possible to this problem:

(a) Specification of mechanisms whereby technical progress can be expected to be systematically Harrod-neutral. This approach has been discussed by a number of writers[16] and Drandakis and Phelps have argued that 'Bowley's Law'[17], which claims that factor-shares have remained constant for a very long period, demands that such a mechanism be found. It is clear that the apparent constancy of relative shares, which Keynes referred to as 'a bit of a miracle' (141), appears to be inconsistent with non-Harrod-neutral technical progress.

(b) Atkinson (15) has pointed out that steady growth, approximately constant relative shares, non-Harrod-neutral technical progress and an elasticity or substitution not equal to unity are not necessarily inconsistent if account is taken of the *time* that might elapse before it could be perceived that one relative share was tending to zero. He draws 'the tentative conclusion that . . . with quite reasonable values of the parameters the approach to a long-run equilibrium with one factor share zero may take a relatively long time' ((15) p. 143)—in one of his numerical examples the gross share of capital takes 132 years to reach half its initial value. Thus, if the length of the long run is taken into account, it may not be necessary to worry about technical progress not being Harrod-neutral in the short run—but the asymptotic results of the simple growth models would have lost much of their usefulness.

(c) It might reasonably be argued that the necessity of the special case of Harrod-neutrality if steady-growth is to be possible does no more than highlight the weaknesses of models of economic growth dominated by conceptions of steady growth. Hahn and Matthews are unequivocal:

> 'This (the necessity that technical progress be exactly Harrod-neutral) is a further limitation on the value of steady-state growth models as representations of reality' ((85) p. 53).

If we are prepared to assume or assert that technical progress *does* systematically take the special Harrod-neutral form then there is very little difficulty in integrating it into the Harrod or neoclassical models discussed in Chapters 3 and 4. We have already noted that Harrod-neutral or labour-augmenting technical progress is *exactly* equivalent to an

[15] See, for example, Atkinson (15) pp. 137–8 or Burmeister and Dobell (34) pp. 79–80.
[16] e.g. Kennedy (135), Drandakis and Phelps (62), Samuelson (222) and Ahmad (3). See Section 8.3(ii).
[17] A. L. Bowley wrote a series of papers at the end of the nineteenth and the first decades of the twentieth century which suggested that the relative shares of capital and labour in a number of countries had been relatively constant over a long period. See Deane and Cole (51) Ch. VII.

increase in the size of the labour force. A thorough discussion of Harrod or neoclassical models of economic growth incorporating technical progress is therefore not really necessary—the constant rate of growth of the labour force, n, in Chapters 3 and 4 can be simply reinterpreted to include the effects of Harrod-neutral technical progress. Nevertheless, we briefly discuss the effects of the inclusion of technical progress in the Harrod model and examine its implications for the neoclassical model.

(i) The Harrod Model

In Chapter 3 we saw that if steady growth at full-employment were to be possible in a Harrod-style model, then the warranted rate of growth, s/v, would have to equal the natural rate of growth which was equal to the rate of growth of the labour force, n. If Harrod-neutral technical progress is proceeding at an exogenous constant proportional rate of m, the natural rate of growth is redefined to equal *the sum* of the constant rate of growth of the labour force, n, with the rate of technical progress—the natural rate of growth $= n + m$. In this situation, the 'effective' labour force, $A(t)L$, grows for two reasons.

(1) the actual number of men is increasing at a constant proportional rate of n and

(2) the 'efficiency' of the men in producing output is increasing at a constant proportional rate of m. In the words of Hahn and Matthews:

> 'Population growth causes there to be two men where there was previously one; Harrod-neutral technical progress causes one man to be able to do twice what he could have done previously' ((85) p. 50).

With the inclusion of Harrod-neutral technical progress the necessary condition for steady growth with full employment becomes

$$s/v = n + m \tag{7.4.6}$$

What we have described as the First and Second Harrod Problems (see 3.4) are unaffected by the inclusion in the model of Harrod-neutral technical progress. The parameters s, v, n, and m are all independently determined and there is still no reason whatsoever why equation (7.4.6) should be satisfied. Since the Harrod stability problem is independent of the natural rate of growth it is clear that it cannot be affected by the inclusion of technical progress. Thus, the inclusion in the Harrod model of technical progress, even in its special Harrod-neutral form, does not substantially alter the vision of growth with which he has come to be associated.

(ii) The Neoclassical Model of Growth

The neoclassical model of economic growth discussed in Chapter 4 is not altered a great deal by the inclusion of technical progress. Retaining Assumptions 4.2.1–5 but replacing the aggregate production function of equation (4.2.6) by

$$Y = F(K, A(t)L) \tag{7.4.7}$$

as the formal representation of a technology incorporating labour-augmenting or Harrod-neutral technical progress, the remainder of the analysis of the model is almost identical to that of Chapter 4 except that the labour force measured in natural units (workers) is replaced by a labour force measured in efficiency units—the 'effective labour force' $A(t)L$. If we write $\bar{y} = Y/A(t)L$, output per worker in efficiency units, $\bar{k} = K/A(t)L$, capital per worker in efficiency units and assume that the Harrod-neutral technical progress proceeds at the constant proportional rate m (i.e. $\dot{A}/A = m$) then the analysis of equations (4.3.1)–(4.3.11) can be simply repeated with all per-worker magnitudes now being measured in efficiency units and the fundamental equation of neoclassical economic growth will emerge[18] as

$$\dot{\bar{k}} = sf(\bar{k}) - (n + m)\bar{k} \tag{7.4.8}$$

In Fig. 7.6, the vertical axis measures output per worker measured in efficiency units while the horizontal axis measures the capital–labour ratio

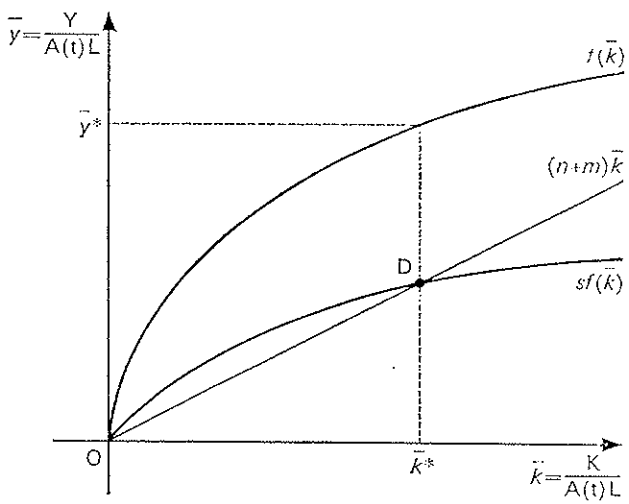

Fig. 7.6

[18] The student would be well advised to derive the equation for himself along the lines indicated in the text.

in efficiency units. The curve $sf(\bar{k})$ represents savings per 'effective' worker and steady-state growth occurs at the point D where this curve intersects the line $(n + m)\bar{k}$. In totally analogous fashion to our analysis of Fig. 4.1 we can show that if $\bar{k}$ is less than $\bar{k}^*$ then $\bar{k}$ is increasing and, conversely, if $\bar{k}$ is more than $\bar{k}^*$, $\bar{k}$ is falling. Thus, given the revised fundamental equation of neoclassical economic growth, (7.4.8), we can see that the economy gravitates smoothly to a constant level of the capital–labour ratio in efficiency units, $\bar{k}^* = K/A(t)L$ and a constant level of output per worker in efficiency units, $\bar{y}^* = Y/A(t)L$. If $K/A(t)L$ is constant then the capital stock must be growing at the same rate as the effective labour force, $A(t)L$. The rate of growth of the *effective* labour force is equal to the sum of the rate of growth of the number of workers, n, and the rate of Harrod-neutral technical progress, m. Thus, the long-run rate of growth of the capital stock is the natural rate of $n + m$. Similarly, if $Y/A(t)L$ is to remain constant then the rate of growth of output must equal the rate of growth of the effective labour force, $n + m$. It is important to notice that, although the capital–labour ratio and output per worker in *efficiency units* are constant in steady growth, the actual capital–labour ratio, K/L, and the actual output per worker, Y/L, both grow at a constant proportional rate equal to the rate of labour-augmenting technical progress. Both Y and K have a long-run rate of growth of $n + m$ whereas the actual labour force only grows at the rate n. Since the rate of growth of output per worker is given by

$$\dot{y}/y = \dot{Y}/Y - \dot{L}/L$$

$$= (n + m) - n$$

and that of capital per worker by

$$\dot{k}/k = \dot{K}/K - \dot{L}/L$$

$$= (n + m) - n$$

it is easy to see that Y/L and K/L grow at the rate m in a neoclassical model of growth incorporating technical progress.

In the conclusion to Chapter 4 we noted that the simplest neoclassical model was inconsistent with Kaldor's 'stylized facts' in that it predicted a constant level of output per worker and capital per worker. We note that the introduction of a simple form of technical progress goes a long way towards the removal of this problem in that the measured capital–labour ratio and output per worker are now seen to grow at a constant rate. The results of including technical progress in a simple neoclassical growth model are summarized in Table 7.1[19] Thus, the inclusion of a simple representation of technical progress in the one-sector model of neoclassical economic growth does not substantially alter the harmonious vision but does produce conclusions which are more in accord with what appear to be the stylized facts of economic reality.

[19] All the results which are included in the table but which are not explicitly discussed in the text can be easily derived by a consideration of Fig. 7.6.

Table 7.1

Harrod-neutral technical progress in a simple neoclassical model

Symbol	Meaning	Steady-state Solution
Y	National output	Grows at a constant rate of $n + m$
K	Capital stock	Grows at a constant rate of $n + m$
L	The labour force in natural units	Grows at a constant rate of n
$A(t)L$	The labour force in efficiency units	Grows at a constant rate of $n + m$
K/L	The capital–labour ratio in natural units	Grows at a constant rate of m
Y/L	Output per worker in natural units	Grows at a constant rate of m
K/Y	The capital–output ratio	Remains constant
$Y/A(t)L$	Output per worker in efficiency units	Remains constant
$K/A(t)L$	Capital–labour ratio in efficiency units	Remains constant
C/L	Consumption per worker	Grows at a constant rate of m

7.5 The Measurement of Technical Progress

Since the middle of the 1950s a series of papers, books and monographs[20] has appeared in which, with more or less sophistication, attempts have been

[20] Interest originated with Abramowitz's contribution (1). See also Solow (245), Kendrick (133) and Denison (52) (54).

made to measure the contribution of technical progress to economic growth in a variety of countries. In 1956 Abramowitz had found that almost none of the entire increase in per-capita output in the United States after 1870 could be accounted for in terms of increases in the stock of physical capital or the supply of labour services. Subsequent investigation seemed to confirm that, for most countries,[21] some factor, quickly dubbed 'The Residual', other than the accumulation of capital and labour was responsible for a very high percentage of observed economic growth. The methods by which these conclusions were obtained can be illustrated by a consideration of Solow's approach in his 1957 paper (245). The central theoretical basis of Solow's paper was the derivation of an equation identical to equation (7.4.5) of Section 7.4. He was anxious to emphasize that he 'would not try to justify' his argument 'by calling on fancy theorems on aggregation and index numbers' and commented that 'either this kind of aggregate economics appeals or it doesn't' ((245) p. 312). The central equation of Solow's paper is of the form

$$\dot{y}/y = \dot{A}/A + \Pi_K \dot{k}/k \qquad (7.5.1)$$

which is formally identical to that of equation (7.4.5) and is derived by almost identical means[22]. The rate of growth of output per head, $\dot{y}/y$, is to be equal to the sum of the rate of growth of capital per worker, $\dot{k}/k$, multiplied by the relative share of capital, Π_K, with the rate of technical progress, $\dot{A}/A$. Solow explicitly uses the idea of technical progress 'as a shorthand expression for *any kind of shift* in the production function. Thus slowdowns, speedups, improvements in the education of the labour force, and all sorts of things will appear as "technical change" ' (Solow (245) p. 312). Given equation (7.5.1), an estimate of the rate of 'technical progress', $\dot{A}/A$, can be easily obtained using estimates of the rate of growth of output per worker, $\dot{y}/y$, the rate of growth of capital per worker, $\dot{k}/k$ and the relative share of capital, Π_K. Thus, Solow's method involves treating statistics of the capital share, the rate of growth of output per worker and capital per worker derived from a real world of heterogeneous outputs and items of capital equipment *as if* they were equivalent to the aggregative concepts of the simple neoclassical model[23]. Solow concluded that 90% of the growth in US output per worker during the period 1909–49 was due to the effect of the residual factor, A, which was supposed to measure the effect of technical progress. Many writers found this surprising

[21] Australia appears to be an exception in the post-war period. See the reference to Sampson's work in Harcourt (94) p. 48.

[22] Solow explicitly assumes that technical progress is Hicks-neutral and the aggregate production function can therefore be written in the equally capital and labour-augmenting form: $Y = A(t)F(K,L)$. Equation (7.5.1) corresponds to equation 2a of Solow's paper with the notation slightly changed.

[23] For an interesting simple discussion of Solow's methodology see Harcourt (94) pp. 49–51.

and few did not consider it faintly disconcerting. Abramowitz was, perhaps, the most succinct:

> 'This result is surprising in the lopsided importance which it appears to give to (technical progress or the Residual) and it should be, in a sense, sobering, if not discouraging, to students of economic growth. Since we know little about the causes of (technical progress or the Residual) the indicated importance of this element may be taken to be some sort of measure of our ignorance about the causes of economic growth' ((1) p. 11).

As we will see in Chapter 8 more recent research employing more sophisticated conceptions of technical progress and being more careful with aggregation difficulties associated with the available crude data, has substantially weakened this kind of conclusion to the extent that one prominent and controversial paper by Jorgenson and Griliches (121) has suggested that almost *all* of the observed economic growth in the United States between 1945 and 1965 can be accounted for in terms of the growth of the conventional inputs of capital and labour!

Attempts at measuring the contribution of technical progress to economic growth are returned to in Chapter 8. For the moment we note that, at least for a time, the results of Solow, Abramowitz and others were extremely influential in turning attention towards the causes of the 'Residual' or the rate of technical progress. Few writers on technical progress[24] fail to preface their work with comments suggesting that their interest was first stimulated by the results which apparently demonstrated the pre-eminence of technical progress as a cause of economic growth.

7.6 Weaknesses of the Simple Representation of Technical Progress

The conception of technical progress employed in this chapter is, of course, open to very severe criticism as a theoretical construct and as an attempt at a representation of economic reality. We list some of these criticisms.

(a) *The simple conception of technical progress, as represented by a term which is 'plugged-in' to an aggregate production function, includes* all *factors which serve to shift the production function and excludes some important aspects of real technical progress.*

We have already noted that *any influence which serves to produce a shift in the production function.* is classified as technical progress given the simple ideas used in this chapter. It may, however, be appropriate to distinguish between the different factors that might cause more output to be produced with the same amounts of the inputs or the same amount of output to be

[24] See, for example, Mansfield (171) or Brown (31).

produced with less of one or more of the inputs. Consider, for example, a typist[25] who currently types 2,000 words an hour. If her 'output' increases to 3,000 words an hour the discussion of this chapter would indicate that this increase in output would be attributed to 'technical progress'. But the increase in output per hour of labour could have arisen for a number of fundamentally different reasons:

(i) her typewriter has been replaced by a better machine—perhaps an electric model

(ii) her speed has improved with experience or she has been practising in her spare time

(iii) the keyboard of her machine has been altered[26].

Thus, although the effect is the same the cause may have been different and, in particular, we notice that (i) involves a new item of capital equipment presumably paid for by the employer, (ii) the use of some of the worker's leisure and (iii) the adaptation of an existing machine again presumably paid for by the employer. It is clear that the use of the term technical progress as a 'catch-all' for a variety of different effects could obscure some important features of the economic process.

We have also noted that the simple conception of technical progress in a one-good model of economic growth *excludes* one of the central features of the definition of technical progress specified in Section 7.1. To quote Mrs Robinson:

> 'actual technical change largely consists in altering the nature of consumption goods (by substituting one kind for another—say cotton for linen or rayon for cotton, or by introducing new kinds of goods such as motor cars or television sets)' ((209) p. 65).

It is therefore clear that the conception of technical progress used in this chapter is not a very adequate representation of reality although it could, of course, be argued that this is not too serious a deficiency in an economic *theory*. (See Ch. 1.)

(b) *Technical progress as discussed in this chapter, 'comes from nowhere'. It is completely independent of the rate of capital accumulation and of any other variable in the economic system.*

[25] Professor Solow uses the example of a typist on a number of occasions in his Radcliffe Lectures (252).

[26] When the original Remington typewriter was first introduced the keyboard was apparently specifically designed to prevent the operator from typing too quickly and jamming the relatively primitive mechanism. This keyboard design is now universal although other systems have been designed which substantially increase the speed of a typist with little practice.

We have already noted that technical progress, as discussed in this chapter, 'falls like manna from heaven' and, continuing the religious imagery, Harcourt suggests that:

> 'this disembodied neutral technical progress may be likened to a mysterious manifestation of grace—when two or more, in this case capital and labour, are gathered together in this life, there immediately occurs a rise (of considerable dimensions) in total factor productivity' ((94) p. 48).

Technical progress of this kind is referred to as 'disembodied' to distinguish it from a more sophisticated conception in which new ideas and methods are 'embodied' in new machines (see 8.2). Kaldor has been a particularly strident critic of the idea that technical progress proceeds at an exogenous rate and can be conceptually separated from the effect of capital accumulation:

> 'The use of more capital per worker . . . inevitably entails the introduction of superior techniques which require "inventiveness' of some kind. . . . On the other hand, most, though not all, technical innovations require the use of more capital per man—more elaborate equipment and/or more mechanical power. . . . It follows that any sharp or clear-cut distinction between the movement *along* a "production function" with a given state of knowledge and a *shift* in the "production function" caused by a change in the state of knowledge is arbitrary and artificial' ((128) pp. 595 and 596.).

Ideas of 'embodied' endogenous technical progress together with Kaldor's personal approach to these problems are discussed in Chapter 8.

(c) *Technical progress is costless*

The most obvious characteristic of 'manna from heaven' for the Children of Israel was that, apart from the effort expended in gathering it, it was *free*. It could be argued that the most serious deficiency of our simple conception of technical progress is exactly this point. No resources need be used in discovering new ideas or putting them into practice such that their benefits can be enjoyed. Technical progress appears to provide something for nothing—no private or social cost is incurred in obtaining it. Micro-economic research into the causes of technical progress would emphasize the vast expenditures on research and development which can be interpreted as an investment in obtaining technical progress by the individual firm or industry. We return to this question in Chapter 8.

(d) *For steady growth to be possible technical progress must take the special Harrod-neutral form*

We have already discussed this weakness at some length and have noted that it could be argued that it represents a weakness of the conventional method of analysing the process of growth, the steady state, rather than a weakness of the particular representation of technical progress.

In this chapter we have developed a simple representation of technical progress, examined the conventional methods whereby its bias can be classified, and investigated the implications of its injection into the simple models of economic growth of Chapters 3–5. It seems clear that, for the host of reasons outlined in the chapter, this conception of disembodied technical progress proceeding at an exogenous constant proportional rate can hardly be thought of as particularly satisfactory. Attempts at devising a more sophisticated conception constitute the subject matter of the next chapter.

chapter 8
The Transmission and Causation of Technical Progress

8.1 Introduction

The representation of the effects of technological progress introduced in Chapter 7 had the undoubted merit of simplicity. When modified by its inclusion, the simple Harrodian and neoclassical models of economic growth were both better able to conform to the 'stylized facts' of growing economies. The bias of its effects could be neatly classified by examining the change in the ratio of the relative shares in national income of capital and labour—along paths of growth involving a constant capital–labour ratio (Hicks) or a constant capital–output ratio (Harrod). Moreover, its inclusion in a neoclassical aggregate production function provided a very convenient, if controversial, method of isolating its contribution to the economic growth of a real economy (7.5). However, as was emphasized in Section 7.6, this simple conception of technical progress is subject to severe criticism if it is meant to represent the sum total of the effects of increases in 'the social pool of knowledge of the industrial arts' (Schmookler (234 p. 1). The convenience of simplicity cannot obscure the dangers of over-simplification. The technological progress of Chapter 7 'floats down from the outside' (Solow (247) p. 90), is exogenous to the economic system and is costless. It is not surprising that many writers have attempted to develop and analyse more sophisticated conceptions of technical progress which take into account the objections to a simpler treatment. Some of these approaches are discussed in this chapter. The emphasis here is on the statement of central ideas rather than on systematic exposition partly because of space constraints and partly because a complete treatment of many of the newer theories would be beyond the scope of an introductory volume. Section 8.2 discusses the means whereby technological progress can be 'embodied' in new capital equipment and hence transmitted into the economy as technical progress. Section 8.3 introduces some of the ways in which technical progress can be conceived of as endogenous to the economic system. The footnotes and structured reading are intended to provide an avenue to the advanced literature for the enthusiast.

8.2 The Transmission of Technical Progress

The technical progress of Chapter 7 was likened to 'manna from heaven'